UNDERSTANDING CHILDREN'S CAUSAL REASONING
DURING COLLABORATIVE DISCUSSIONS

BY

SHUFENG MA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Educational Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Professor Richard C. Anderson, Chair
Professor John E. Hummel
Associate Professor Jennifer Cromley
Assistant Professor Emma Mercier

ABSTRACT

This study aims to understand the construction of multilink causal reasoning chains during collaborative discussions in elementary school classrooms. The construction of reasoning chains was investigated in 24 collaborative discussions involving 160 underserved fifth-grade children. The effects of group features, individual characteristics, and moment-by-moment situational influences on seven causal chain models were tracked in the discussions. Results indicated that students who were more talkative, had better oral English, and were more liked by their classmates were more likely to produce causal chains. The turn-by-turn analysis of chain construction revealed that once a causal chain was initiated, it was likely to continue for at least three speaking turns. Leaders and socially centered students supported other group members, the shy and quiet students, to extend chains of reasoning. Agreement among group members and support from leaders and socially centered students extended the chain of reasoning. However, refutation and disagreement stopped the chain because the group had to resolve disputed ideas in order to develop a shared understanding. A temporal analysis of chain production indicated that chain construction speeds up over the course of a discussion. Students who produced more causal chains during the discussion also generated more causal chains in an individually written essay after the discussion. A causal analysis showed that peer dialogue mediates the effects of social and cognitive characteristics on chain production in the essay and the mediating effect increases with increases in the number of peer-generated causal chains during the discussions. Overall, this analysis of the social construction of multilink causal reasoning chains provides distinctive new evidence that enabling meaningful interaction among children promotes their higher-level cognitive development.

To my sister and my parents

for their immeasurable support

ACKNOWLEDGEMENTS

It has been an incredible and life-changing journey to study at the University of Illinois. At every point along the way, I have received tremendous help and support from many people. I am hugely indebted to my advisor Prof. Richard Anderson for his mentoring. Beyond everything, I want to thank him for having faith in my potential and choosing me as his student. I would also like to thank him for all the guidance he gave me when I felt lost, for all the support he gave me when I was uncertain and worried, for all the encouragements he gave me during my downtime, for all the jokes he told during our meetings, for all the wisdom and life experiences he shared with me, for being extremely flexible with our meeting time, for all the rides to Jimmy John's, and most importantly, for being my role model and showing me what it is like to be a passionate researcher, a generous collaborator, a humble scholar, a good writer, and overall a true intellectual.

I am highly indebted to and thoroughly grateful to Prof. Jennifer Cromley for her immense interest in my topic of research, for her insightful and honest advice on my research, teaching, job search, and career development, for taking a great amount of time to read my papers and write reference letters for me, for giving me the opportunity to learn how to review a journal article and at the same time how to pet a dog, for being so responsive and approachable when I needed help and, last but not least, for showing me what a great teacher looks like.

I am also very grateful to Prof. Emma Mercier for her support at several important steps of my Ph.D. life. I want to thank her for enhancing my understanding of collaborative learning through well-explained knowledge construction theories and well-structured collaborative activities in her wonderful class. I am also very grateful to her for reading my papers and for having the privilege to receive very constructive feedback on my writing.

I am grateful to Prof. John Hummel for being so kind as to show interest in my research and for sharing his work on role-based relational thinking, which inspired me to start a line of research on multi-link causal reasoning. I would also like to thank him for the time and energy he spent on reading my dissertation and the invaluable advice that helped me improve this study.

I want to express my gratitude and regards to the distinguished faculty members in the College of Education—Prof. Liz Stine-Morrow, Prof. Michelle Perry, Prof. David Brown, Prof. Dan Morrow, Prof. Kiel Christianson, Prof. Robb Lindgren, Prof. Carolyn Anderson, Prof. David Zola, Prof. Jinming Zhang, Prof. Hua-Hua Chang, Prof. Daniel Walsh and Prof. James Rounds, for all their guidance and support throughout the process of helping me become the best teacher and researcher I can be. Special thanks to Prof. Hans-Friedrich Köhn (Frieder) for teaching me introductory statistics for two semesters, which built a strong foundation for me to learn more advanced statistical methods in the following years. His continued passion on research after two years fighting with leukemia has really inspired me to embrace every moment of my life.

I would not enjoy my life here in Champaign-Urbana without my wonderful colleagues and friends. My deepest gratitude is extended to Matthew Sutton, dearest Miss V., Bemi Ekwejunor, Chan Yu, Emilie Palacios, Jingjing Sun, Liang Tao, Xiao Yang, Yuxin Deng, Shuo Xu, Tzu-Jung Lin, Jie Zhang, Judy Hsu, and many more, for their care, love, support, company, and life-long friendship.

This journey would not have been possible without the support of my family. I am hugely indebted to my twin sister Yefeng Ma for her unconditional love and her unceasing effort to teach me the importance of having a balanced life. I am especially grateful to my parents for encouraging me in all of my pursuits and inspiring me to follow my dreams, and for always being ready to answer my phone calls. I am so lucky and blessed to have these people in my life.

TABLE OF CONTENTS

LIST OF TABLES..viii

LIST OF FIGURES..ix

CHAPTER 1: INTRODUCTION..1

CHAPTER 2: LITERATURE REVIEW..4

2.1 Theories of Causal Reasoning..4

2.2 Collaborative Reasoning Discussion..10

2.3 Peer Collaboration and Shared Understanding...14

2.4 Multilink Causal Reasoning...20

2.5 The Rationale for the Present Study..24

CHAPTER 3: METHOD...33

3.1 Design and Procedure..33

3.2 Participants..34

3.3 Discussion Transcripts...37

3.4 Assessments..41

3.4.1 Pretest measures..41

3.4.2 Posttest measure..43

3.5 Coding Multilink Causal Reasoning in Collaborative Discussions.....................44

3.5.1 Identifying patterns of multilink causal reasoning.......................................44

3.5.2 Identifying structures of multilink causal chains...49

3.5.3 Coding features of multilink causal links..53

3.5.4 Coding elements of multilink causal links...54

3.6 Coding Discourse Moves....................57
3.6.1 Coding extending moves....................57
3.6.2 Coding leadership moves....................59
3.6.3 Coding reasoning moves....................60
CHAPTER 4: RESULTS....................64
4.1 Pretest Performance....................64
4.2 Analyses of Features and Structures of Multilink Causal Reasoning....................66
4.2.1 Descriptive statistics of chain production....................66
4.2.2 The influence of group features on chain production....................67
4.2.3 The effect of individual characteristics on production of causal chains....................69
4.2.4 Analyzing the length of multilink causal chains....................75
4.3 Temporal Development of Multilink Causal Reasoning....................76
4.4 Statistical Discourse Analyses of Turn-by-Turn Chain Production....................79
4.4.1 The proximal effects of extending moves....................79
4.4.2 The proximal effects of leadership moves....................84
4.4.3 The proximal effects of reasoning moves....................87
4.5 A Causal Analysis of Peer Modeling and Individual Outcome....................89
CHAPTER 5: DISCUSSION....................97
5.1 Major Findings on the Development of Multilink Causal Reasoning....................98
5.2 The Momentary Social Process in which Multilink Causal Reasoning Emerges......105
5.3 Some Limitations of this Study....................120
5.4 Conclusions and Implications....................122
REFERENCES....................125

LIST OF TABLES

Table Page

1 Number of Children and Duration of Talk in Collaborative Groups................................37

2 Distribution of Speaking Turns in Each Discussion Group...41

3 Decision Tree of Extending Moves..57

4 The Coding Scheme of Leadership Moves...59

5 The Coding Scheme of Reasoning Moves...62

6 Correlations of Seven Social Characteristics..65

7 Spearman Partial Correlations of Group-Level Social and Cognitive Features and Measures of Multilink Causal Reasoning..68

8 Means and Standard Deviations of Measures of Multi-Link Causal Reasoning by Gender, Ethnicity, and Expert Groups 71

9 Proximal Effects of Previous Speakers' Characteristics and Extending Moves...............83

10 Proximal Effects of Leadership Moves on Multilink Causal Reasoning..........................86

11 Proximal Effects of Reasoning Moves on Multilink Causal Reasoning..........................88

LIST OF FIGURES

Figure Page

1 Use of academic vocabulary over six intervention sessions..36

2 The multilink reasoning structure of ecosystem-economy model (Model A) 45

3 The multilink reasoning structure of the ecosystem-oxygen model (Model B)..................46

4 The multilink reasoning structure of the leftover model (Model C) 47

5 The multilink reasoning structure of competition model (Model D) 47

6 The multilink reasoning structure of the compensation model (Model E) 48

7 The multilink reasoning structure of the supply model (Model F) 48

8 The multilink reasoning structure of the tourism model (Model G) 49

9 Average time interval between successive pairs of causal chains.......................................77

10 The distribution of five types of leadership move among students and teachers 84

11 The Poisson model for analyzing the effect of individual characteristics on children's chain production in the wolf letter..91

12 The path model for analyzing the effect of self-initiated links during discussion on children's chain production in the wolf letter...92

13 The path model for analyzing the effect of peer-initiated links during discussion on children's chain production in the wolf letter...93

14 The path model for analyzing the interaction effect of self-initiated links and peer-initiated links on children's chain production in the wolf letter...94

15 The path model for analyzing the interaction effect of peer-initiated links and talkativeness on children's chain production in the wolf letter...96

CHAPTER 1

INTRODUCTION

My study builds on the existing research on how causal reasoning develops in children. A large body of research has documented children's use of causal inferences to interpret the outside world (e.g., Callanan & Oakes, 1992; Hickling & Wellman, 2001; Inagaki & Hatano, 2006; Legare, Schult, Impola, & Souza, 2016; Mogar, 1960). Causal reasoning is the higher-order cognitive skill by which people form abstract representations and create complex causal structures (Sobel & Legare, 2014). Previous research has produced several theories to explain how causal relations are formed (see reviews such as Khemlani, Barbey, & Johnson-Laird, 2014 and Sobel & Legare, 2014). Two popular theories posit that causal reasoning is either driven by learned knowledge of causal relationships (i.e., causal relations are deterministic; Kant, 1781/1997; see also Khemlani et al., 2014) or is a natural tendency to connect events according to structural resemblance or temporal and spatial contiguity (i.e., causal relations are probabilistic; Hume, 1748/1999; see also Hastie, 2016).

If we take a developmental perspective, the two accounts can be integrated. Children's causal learning in the early stages of life is "a form of association of experiences" (Sobel & Legare, 2014, p. 418) based on constantly updating domain-specific knowledge of temporal priority, spatial contiguity, and patterns of covariation (Hastie, 2016), and later develops into complex, multilevel causal models or general principles that can be applied to substantive cases of causal reasoning (Khemlani et al., 2014). Seeing the two stages as part of one continuous process explains the development of children's causal reasoning. However, the how and why of the construction of causal models is still being explored.

So far most research on children's causal reasoning has been conducted in well-controlled, simplified psychological laboratory settings and has been focused on the learning mechanism and developmental pattern; however, not much is known about children's productive use of causal reasoning in natural social settings. My dissertation extends previous research by investigating a child's ability to generate a chain of causal reasoning during collaborative discussions. In natural social settings, such as a collaborative discussion, the complex structure of causal reasoning makes high demands on cognitive processing and requires the ability to discern deep structures despite dissimilarities in surface features (Chi & VanLehn, 2012). Children, especially in underserved communities, often lack experience in connecting ideas, especially when making connections that take several reasoning steps (Flood, 2010) because language art classes in schools usually focus on children's basic language skills such as decoding, pronunciation, spelling, vocabulary, grammar, and reading comprehension (Heilig, 2011; Menken, 2009). Children often struggle with organizing components within a system and tend to oversimplify complex relations (Assaraf & Orion, 2010; Barman & Mayer, 1994).

By interacting with others, children gain experience with connecting ideas. A large and growing body of research has shown that well-designed and intellectually-stimulating student-student or student-teacher interaction improves aspects of children's higher-order thinking (e.g., Barron, 2000; Kuhn, Shaw, & Felton, 1997; Osborne, 2010; Webb & Mastergeorge, 2003). Through social interaction, learners are able to solve problems with the help of a more competent person when not able to do so on their own (Vygotsky, 1978). Teachers are usually considered to be the ones who provide such support; however, research has also documented the value of assistance that competent peers provide to their less proficient classmates (e.g., Azmitia, 1988; Fawcett & Garton 2005; Garton & Pratt, 2001).

This study investigates children's productive use of causal reasoning in an interactive learning environment. Special attention is devoted in this study to classrooms that utilize the

Collaborative Reasoning approach (CR; Anderson, Chinn, Waggoner, & Nguyen-Jahiel, 1998). CR is a highly interactive method that has shown promise in improving cognitive skills, including argumentative skills (Anderson et al, 2001; Reznitskaya, 2009), language skills (Zhang, Anderson, & Nguyen-Jahiel, 2013), analogical reasoning (Lin et al., 2012), and relational thinking (Lin, Anderson et al., 2015). In Collaborative Reasoning discussions, students work in small groups discussing controversial issues with considerable autonomy and freedom. They are instructed to provide reasons and evidence for their positions and encouraged to challenge each other's opinions. My colleagues and I have found that children who had six weeks of experience with Collaborative Reasoning and other forms of collaborative group work produced longer chains of reasoning (many 4-7 link chains) in oral narratives (Ma, Anderson et al., 2017), written policy-decision essays (Lin et al., 2011), and argumentative oral interviews (Ma & Anderson, 2015), in comparison with children who had six weeks of teacher-led direct instruction (many 1-3 link chains).

This study builds on previous work and identifies the ways that causal reasoning in the form of multilink causal chains is constructed in group discussions. The premise of this study is that children's causal reasoning benefits from quality social interaction with peers during collaborative discussion, as it creates an open and engaging environment for learners to provide explanations, elaborate and support ideas with evidence, consider alternative explanations, and challenge one another's ideas. Such activities draw on the ability to identify relationships and make connections, which is the basic operation of causal reasoning. This study is expected to enable a better understanding of the relationship between social interaction and cognitive development through sophisticated analysis of the pattern of connecting reasoning links within a speaking turn, within a discussion, within an individual child, and within a group of children.

CHAPTER 2

LITERATURE REVIEW

This literature review is composed of five sections. The first section summarizes relevant theories of children's causal reasoning. The second section introduces the Collaborative Reasoning approach, which is an interactive learning format that enables meaningful peer interaction and provides social context for evaluating children's causal reasoning. The third section reviews the literature on peer collaboration and the assessment of shared understanding in collaborative group work. The fourth section introduces one specific type of causal reasoning, multilink causal reasoning. The fifth section lays out the rationale for this study and presents the research questions.

2.1 Theories of Causal Reasoning

In the past four decades, a large body of research has documented children's use of causal reasoning in physical, social, and psychological domains (see a detailed review in Gopnik & Wellman, 2012). Children make causal inferences to interpret the outside world. For example, Inagaki and Hatano (2006) found that preschoolers developed the understanding of 'vitalistic causality' to explain why people need food and water to exist. Children's causal reasoning is not only substantiated by the identification of specific causal relationships, but also by the construction of abstract frameworks or generalizable knowledge. Vosniadou and Brewer (1992) found that elementary school children, when challenged to explain the shape of the earth, collectively formed six different mental models based on interpretations of everyday experience.

Previous research has produced several theories to explain how causal relations are formed (see reviews such as Khemlani, Barbey, & Johnson-Laird, 2014 and Sobel & Legare, 2014). Two popular theories posit that causal reasoning is either driven by learned knowledge of causal relationships (i.e., causal relations are deterministic; Kant, 1781/1997; see also Khemlani, Barbey, & Johnson-Laird, 2014) or is a natural tendency to connect events according to structural resemblance or temporal and spatial contiguity (i.e., causal relations are probabilistic; Hume, 1748/1999; see also Hastie, 2016).

The first theory is drawn from the Kantian view of causation, which argues that children's causal reasoning is built on the identification of causal structures or abstract domain-general knowledge, which is used to perceive higher-level relationships between objects or events. Kant (1781/1997) considers causation or causal relationships to be innate assumptions, by which people provide reasonable and logical explanations for observed phenomena. Causes and effects are independent entities, and causal reasoning is the process of connecting causes with effects under the guidance of some general principles. The understanding of such general principles is necessary to perform causal reasoning; for example, the general principle that gravity causes an object to fall explains the fact of an apple falling to the ground. According to the Kantian view of causation, when a child can perform causal reasoning, he or she can develop an understanding of some generalizable patterns (i.e., *schema* or *regularities*) of a series of events or phenomena, and then match the causes with the effects based on these patterns. Children may not know exactly how the mechanism works, but they intuit the existence of such principles or rules.

The second theory of causal reasoning, proposed by David Hume, argues that causal reasoning is not built on general principles, but rather is a tendency to connect objects or events in terms of structural resemblance and temporal or spatial contiguity (Hume, 1748/1999). Hume treats causality as a habit of mind rather than an application of learned principles. When one event is observed to occur right after another event, it is natural to think that the former event causes the second event. Causal reasoning, from the Humean perspective, is regarded as the association of experiences based on temporal priority, contiguity, and covariation (Koslowski & Masnick, 2010). That is to say, a causal relation can be inferred if the change of an event or an object occurs concurrently with, or shortly after, the change of another event or object.

Most psychological research in children's causal reasoning examines children's sensitivity to Humean indices of causality, i.e., temporal priority, contiguity, and covariation (Koslowski & Masnick, 2010). There is an increasing amount of research that focuses on developing probabilistic models (e.g., the causal graphical model; Glymour, 2001; Gopnick et al., 2004; Sobel & Legare, 2014) to explain how children use prior knowledge to draw causal connections for new phenomena based on probabilistic models and the Bayesian learning mechanism (see a detailed review in Griffiths & Tenenbaum, 2009).

The causal graphical model provides a computational perspective on children's causal reasoning in well-controlled laboratory settings. This model is built upon three assumptions (Sobel & Legare, 2014). First, the relationship between two objects or events complies to some formal principles (Humean indices) that are independent of the situation, such as temporal priority, spatial contiguity, or covariation. Second, children's causal reasoning follows the Markov assumption, which says that the occurrence of one event only produces one son event. In other words, if event A leads to the occurrence of event B, and event C leads to the occurrence of event B, when discussing the role of event A on event B, the possible influence of event C on this relation is not taken into account. For example, drinking alcohol can cause insomnia and stress can also cause insomnia, but drinking alcohol will lead to insomnia regardless of whether the person is stressed or not. However, in natural settings people's causal reasoning often violates Markov's assumption because they tend to bring in prior knowledge to examine whether there are other events that could possibly elicit the occurrence of the events under consideration

(Park & Sloman, 2013). The third assumption is that children's causal reasoning is based on relationships that are empirically existent in the world, not based on hypothetical relationships. In other words, when a faithful causal relationship exists between event A and event B, causal graphical models do not account for the coincidence of a hypothetical event C that could possibly boost or prevent the occurrence of event A or event B.

These three assumptions suggest that causal graphical models can explain the most simplified form of causal reasoning, namely the process of how the change of one event or object affects the change of a relative event or object. It is important to note that causal graphical models are domain-specific graphical networks that represent "a list of all possible combinations of events under consideration and the probability that each combination occurs" (Sobel & Legare, 2014, p. 415). In a given situation, a person selects the most likely graphical network to explain the covariation of events or objects based on his or her prior knowledge.

Causal graphical models assume that children's understanding of formal principles (prior knowledge) is well-formed and therefore ready to be applied to substantive cases. Prior knowledge with high relevancy to the problem is more likely to be retrieved than knowledge that appears unrelated. Higher levels of prior knowledge produce more observations about cause-effect relationships. However, in natural settings, children constantly receive information from others, which often leads to the expansion or modification of their knowledge and therefore increases the complexity of the interpretation of cause-effect relationships.

Xu (2007) propose a theory called *rational constructivism* to understand children's development of causal reasoning, which integrates the computational perspective with constructivist's viewpoints. The main idea of rational constructivism is that people first form an intuitive theory about observed phenomena based on statistical inferencing (rational) and gradually update this theory with new evidence (constructivism) obtained from interacting with the world or observing the actions of other people (Gopnik & Wellman, 2012).

The formation of intuitive theory is based on rational statistical inferences—children are sensitive to the relationship between sample and population and they are able to perform statistical computation with new information (Xu & Kushnir, 2013). Their sensitivity to probabilistic relations can further guide them to make predictions, evaluate alternative hypotheses, or assess whether a piece of evidence is sufficient to support the hypothesis. For example, when young children were asked to guess the color of a ping-pong ball after the examiner randomly grabbed a ball from a basket of red and white balls that are mixed in a fixed proportion, the children were more inclined to guess the color that seemed to have a larger amount in the basket (Xu & Denison, 2009). This study showed that participants were aware of the causal relationship between the sample and the population (prior knowledge) and were able to perform causal reasoning in specific cases based on statistical inferences.

The theory of rational constructivism emphasizes individual's development of causal reasoning rather than the individual's role in promoting co-development of causal reasoning in a social context. Although rational constructivism argues that a person modifies his or her intuitive theory of causal relationships based on other people's input, the formation and update of causal understanding is mainly ego-centric. The effect of social interchange is under-evaluated in the theory of rational constructivism, as is mutual influence between learners. It is important to take into account, however, that in natural social settings, knowledge develops as a result of social interaction. Learners consistently update their own causal understanding and influence others' causal reasoning as they take in new information from others, identify cognitive conflicts between their prior knowledge and others' contradictory experiences, and use this information as a basis to restructure their original knowledge and to restore cognitive equilibrium (Piaget, 1952; Golbeck, & El-Moslimany, 2013).

To this point, not much research has documented the way in which peer interaction affects children's causal reasoning specifically, but a large and growing body of research has shown that well-designed and intellectual-stimulating student-student or student-teacher interaction improves aspects of children's higher-order thinking and reasoning, such as argumentation (Chinn & Clark, 2013, Evagorou & Osborne, 2013; Kuhn, Shaw, & Felton, 1997), problem-solving (Barron, 2000; Hmelo-Silver et al, 2013, Phelps & Damon, 1989; Webb & Mastergeorge, 2003; Vye et al, 1997), analogical reasoning (Lin et al, 2012; Mason, 1996), and scientific thinking (Golbeck, 1998; Mason & Santi,1998; Osborne, 2010).

The reason that constructive peer interaction promotes high-level thinking and reasoning ability is that learners have the opportunity to generate more elaborated explanations (Ploetzner, Dillenbourg, Praier, & Traum, 1999). When a person explains an idea to others, plenty of opportunities arise for the learner to modify his or her preexisting knowledge and construct new knowledge. While applying previous knowledge to the new problem, missing knowledge can be identified, new inferences can be drawn from the given information, and certain comprehension gaps or failures can be located. Children are expected to interact with others to refine their understanding of cause-effect relationships when performing causal learning in a social context. While interacting with peers or more experienced adults, children encounter cognitive conflicts or observe inconsistencies in their own thinking. By resolving cognitive conflicts or correcting inconsistencies, children reconstruct or update their mental models for causal relations.

The investigation of children's causal reasoning has hitherto focused on how children identify simplified causal relationships in laboratory settings. To my knowledge, no research has explored children's causal reasoning in natural settings that involve a constant exchange of ideas, a give and take of information, negotiation of meanings, and pursuit of shared understanding. The reasoning process is more complex than that described in causal graphical models because the construction of causal models not only depends on one's own prior knowledge, but is also heavily influenced by the social environment.

It is apparent that the investigation of causal reasoning cannot be separated from its social context. In the present study, children's causal reasoning is evaluated in the collaborative discussion setting. Collaborative discussion provides an authentic social context for students to interact with each other and to co-construct knowledge. The social dynamics of a collaborative discussion make causal reasoning more complex than making an inference in a controlled psychological experiment. For example, the co-constructed causal relationships may build on multiple causes or multiple effects. The relation can be transitive. In this case, event A causes event B and event B causes event C, so that a change in event A also causes a corresponding change in event C. In these situations, a causal chain containing multiple causal relations is formed. To further explain the social context in which student produce causal reasoning, the next section provides a detailed description for one type of collaborative discussion, the Collaborative Reasoning discussion.

2.2 Collaborative Reasoning Discussion

This study examines children's causal reasoning in one specific type of social setting—the Collaborative Reasoning discussion (CR; Anderson et al., 1998). CR was one of the various interventions that have been proposed to increase social interactions in classrooms and to improve both the quantity and quality of classroom talk. CR was first proposed as a literature-based reading program, aimed at promoting children's critical reading and higher-order thinking.

In traditional reading classes, teachers encourage students to take an efferent stance and focus on the interpretation of story characters' behaviors, feelings, and experiences (Rosenblatt, 1989). In contrast, Collaborative Reasoning approach creates an open atmosphere for students to analyze and evaluate story characters' behaviors and predict the possible outcomes to such behavior (Waggoner, Chinn, Yi, & Anderson, 1995). Instead of the teacher guiding students to locate relevant information in the text, students form into groups with a balance of gender, ethnicity, and academic ability to engage in a peer-led, open-format group discussion and deal with a central question that usually contains controversy or a moral dilemma. The central question is formulated by the teacher based on the story. During the discussion, children present their initial position to the whole group, provide reasons to support that position, look for evidence from the text or personal experience to justify the position, and challenge each other's arguments when an assumption is unwarranted or the other side of arguments remains undeveloped (Anderson et al., 1998).

Open participation and the critical/analytic stance are two fundamental features of Collaborative Reasoning. Compared with traditional instruction, in CR discussions students themselves control the whole discussion and are encouraged to freely participate in the discussion and to challenge others' opinions. During the discussion, children learn to use different argument strategies and to effectively utilize their life experience to support their final position on the issue (Clark et al., 2003).

Even though CR is a student-centered instructional approach, the teacher still plays an important role in maintaining an open discussion. In the CR approach, the teacher's task is not to cover all the content knowledge, but to nurture students' thinking and reasoning. A successful CR discussion relies on the collaboration of teacher and students. Teachers are supposed to relinquish control over turn taking, choice of topic, and interpretation of content to the students (Chinn et al., 2001). The students, on the other hand, must take more responsibility for managing the discussion and articulating their thinking. Specifically, teachers' instructional moves include the following aspects (Chinn, Anderson, & Waggoner, 2001):

- Prompt for position, reason, evidence, clarification, alternative perspective, or counterarguments.
- Model the process of making arguments, defending positions, or challenging others' opinions so that students learn how to articulate their thinking.
- Encourage students to express their opinions freely and challenge each other when they disagree.
- Give students the freedom to decide what to say and when to say it.
- Bring up unexplored topics or contradictory viewpoints when the group has reached a consensus or has only considered one side of an argument.
- Sum up what has been said and how the argument has developed so that all the group members maintain a shared understanding.

Because the Collaborative Reasoning approach encourages students to take a critical/analytic stance on textual information, students' argumentative skills significantly improve during CR discussions (Reznitskaya et al., 2009). Children take positions, extract evidence from reading material, articulate reasons and justifications, and make counterarguments and rebuttals. Students establish their own argumentation strategies or imitate other students' argument moves during CR discussions, indicating their development of an argument schema (Reznitskaya, Anderson & Kuo, 2007). Improvements of argumentation skill were documented in similar studies on Chinese and Korean students (Dong, Anderson, Kim, & Li, 2008) and

Spanish-speaking English language learners' learning in CR classrooms (Zhang et al., 2013). In addition, argumentation skill was found to transfer between different contexts and to snowball within groups. Children successfully transferred the argumentation skill they learned in CR discussion to independent essay writing (Dong et al., 2008; Kim et al., 2011; Reznitskaya et al., 2001).

Collaborative Reasoning is also an effective approach to promote the consideration of multiple sides of an issue and to weigh the different arguments. Morris et al (2012) found that students who participated in collaborative discussions were nearly three times more likely to spontaneously consider both sides of a controversial issue than students who received teacher-directed instruction or than uninstructed control students. Zhang et al (2016) examined children's decision making ability in reflective essay writing following participation in CR discussions. The essay writing task was based on a story about a kid named Jack who has to decide if he should tell on his friend Thomas for cheating in a model car competition. The results of this study indicated that students in collaborative groups recognized more than one side of a dilemma, considered a variety of reasons, and weighed the importance of those reasons, as compared to students in direct instruction or uninstructed control condition.

Other studies have documented the development of ideas and strategies in Collaborative Reasoning discussions. An effective argumentative strategy spreads throughout the whole group after its first appearance (Anderson et al., 2001). Also, previously-used analogies are frequently adopted by the group once they are generated, and the introduction of new analogies increases overall (Lin et al., 2012). English language learners take more speaking turns in a CR discussion and their vocabularies expand as they adopt expressions from other students' speech (Zhang et al., 2013). Morris et al (forthcoming) sampled 6 four-minute excerpts from twelve different classrooms for a total of 72 episodes of classroom talk. They searched these transcriptions for low-inference discourse markers of connected and elaborated talk (i.e. but, because, so, if, then, and). From this analysis, they showed that the rate of students' usage of coordinating conjunctions in classroom talk was four times higher among students who participated in collaborative discussion than among students who received direct instruction.

Built on the promising results of Collaborative Reasoning on various aspects of cognitive development, this study aims to understand how *causal reasoning* is nurtured in CR discussions. The guiding theory of this study is that children's causal reasoning benefits from quality social interaction with peers during collaborative discussion, as it creates an open and engaging environment for learners to provide explanations, elaborate and support ideas with evidence, consider alternative explanations, and challenge one another's ideas. Such activities draw on the ability to identify relationships and make connections, which is the basic operation of causal reasoning. To further explain my motivation for studying the development of causal reasoning during collaborative discussion, the next section reviews the literature on peer collaboration and reflects on the assessment of shared cognition during collaborative group work.

2.3 Peer Collaboration and Shared Understanding

According to Vogotsky's sociocultural theory (Vygostsky, 1978), social context shapes the way that learners construct knowledge, manage interaction, and reason. Learners gradually update their understanding of their environment through social interaction and they organize information they receive from other people to develop an increasingly complex mental model (see also Hakkarainen, Paavola, Jangas, & Seitamaa-Hakkarainen, 2013; Palincsar, 1998; Piaget, 1952). Vygotsky (1978) introduced the concept *zone of proximal development* to describe the distance between a learner's actual development level and the potential level of development. He argues that learning starts at the point when children are close to a level of potential development and gradually reach that level through interacting with others. Children's competence increases over time when they interact with a person who provides guidance and support to help them reorganize their current mental structure, especially on tasks that they cannot perform on their own (Vygotsky, 1978).

In collaborative problem-solving tasks, students provide explanations to support each other's comprehension, ask each other questions to check for understanding, and encourage alternative viewpoints from group members (Azmitia, 1996; Hmelo-Silver & DeSimone, 2013). Lu, Chiu and Law (2011) examined how the current speaker's generation of claims, evaluations, and questions affected the next speaker's use of explanations and evidence during online discussions. They found that if a speaker made a new claim or asked a question during the current speaking turn, it was more likely for the next speaker to use that evidence and explanations in the subsequent turn.

Furthermore, Chiu (2008) found that when ninth-grade students worked in groups to solve an algebra problem, disagreement between previous speakers stimulated the current speaker's production of correct ideas. The challenge from peers stimulates learners to revisit their thinking and modify their reasoning in order to develop a new understanding of the problem (Bell, Grossen, & Perret-Clermont, 1985). Schwarz, Neuman and Biezuner (2000) reported the "two-wrongs-make-a-right phenomenon" (p. 461) when they paired up low-achieving tenth and eleventh graders to solve a math problem that children were not able to solve individually. They found that the pair was more likely to produce a right solution when two students disagreed with each other, actively argued about different strategies, and collaboratively tested the hypothesis that each person proposed.

These studies suggest that peer collaboration facilitates the revision of learners' preexisting knowledge structure and fosters joint creation of new knowledge. However, to achieve successful small group work requires more than asking students to work together (Jacobs, 1998). In a well-functioning collaborative group, learners are expected to constantly build and reflect on their own thinking by looking for support from others, filling information gaps, and responding to challenges and disagreements so that the group can maintain a shared understanding of the issue (Roschelle, 1992, Van den Bossche, Gijselaers, Segers, Woltjer, & Kirschner, 2011; Webb, 2013).

The development of a shared understanding is an indicator of the quality of group performance. Barron (2000) compared two sixth-grade triads' performance on a genuine problem-solving task that requires students to work under time and distance constraints while planning a trip for a story character. The results indicated that one of the triad successfully solved all the problems and presented satisfactory answers and sound explanations while the other triad encountered communication and comprehension problems during group discussion and, as a result, was less successful in producing answers and explanations. Three key differences were identified in these triads' interactions: the degree of mutuality; the extent to which students achieved joint focus of attention; and the level of shared alignment or knowledge co-construction. In the less successful group, students generated more self-focused talk, which means they were less likely to build on each other's ideas and were more prone to repeat their own opinions. A good argument was often neglected or rejected without justification. However, in the more successful group, students respected each other's opinion, and clarified and explained rather than simply rejecting or neglecting the other speakers' statements. Each member tried to make sure everyone was on the same page and they used the workbook as the "center of coordination" rather than a "contested territory" (Barron, 2000, p. 430).

The development of shared understanding is more effective when group members build on each other's thinking rather than focusing on their own ideas. If a group member provides an explanation by incorporating other group members' opinions, the explanation is more likely to be accepted or elaborated on by those whose opinions were taken into account. However, if group members all focus on their own thinking, they may each end up with an incomplete and biased solution. In this case, peer collaboration is of minimal benefit to the development of shared understanding in the group. Evagorou and Osborne (2013) compared two dyads' performances when arguing about a socio-scientific issue in an online group discussion. They found that Dyad A produced more claims and evidence, as well as more counter-claims and rebuttals, whereas Dyad B only presented simple claims and directly quoted evidence from the reading material. A detailed analysis of students' discourse indicated that Dyad A asked more questions to clarify the problem, made more explanations to interpret the data, discussed the structure of arguments, and generated analogies to enhance understanding during the discussion. However, Dyad B tended to reach a consensus at the beginning of the task and then each person cited facts located in the text to support the claims. This study reveals that students tend to be content with easy solutions or intuitive explanations if they are not prompted to take complicating facts into consideration. To go beyond intuitive thinking, the group needs to actively engage in the negotiation of meanings by asking each other questions, responding to claims and counterclaims, and discussing the deep structure of knowledge rather than the surface form.

Ellis (2005) summarized Storch (2001) and Wells (1999)'s work on student-student interactions during language learning and concluded that successful small group work provides learners the opportunity to maintain a common understanding through clarification and explanation, to scaffold and expand each other's talk, to co-construct knowledge, and to challenge and reflect on collectively constructed knowledge artifacts. These features may be essential to group success; however, these ideas lack a strong empirical basis. The research on small group work often takes these features for granted (i.e., students maintain a shared understanding and co-construct knowledge when working in groups) and treats them as inherent characteristics when explaining students' interactions in various collaborative group activities (see reviews of theories and empirical studies about collaborative learning in Dillenbourg, Baker, Blaye, & O'Malley, 1995 and O'Donnell & Hmelo-Silver, 2013). However, it is difficult to measure both whether students maintain a shared understanding in small group work and the way in which knowledge is co-constructed among students.

Akkerman and associates (2007) reviewed 22 studies that investigated shared understanding in small group work across the fields of social, educational, and organizational psychology. They found that the concept of shared understanding often lacked clarity and carried multiple meanings in different contexts. An incomplete list of conceptualizations of shared understanding includes shared cognition/understanding (Hatano & Inagaki, 1991; Barron, 2003), common ground (e.g., Stahl, 2005), team mental model (e.g., Mohammed, & Dumville, 2001), and distributed cognition (e.g., Kollar, Fischer, & Slotta, 2007). They clarified and categorized the conceptualizations of shared understanding from two perspectives: the cognitive perspective interprets shared understanding as "cognition at the group-level in terms of similarity, overlap, complementarity, or distribution" (p. 50), whereas the socio-cultural perspective describes shared understanding as "a process of coordination of participants' contributions in joint activity" (p. 50).

The two conceptualizations of shared understanding are assessed in different ways. The cognitive perspective focuses on a learner's construction of mental models and how a learner's mental model overlaps and complements other group members' mental models. The assessment of shared understanding is usually based on an aggregated measure of individual knowledge construction or a collective measure of the group-level product (Mohammed, Klimoski, & Rentsch, 2000). In contrast, the socio-cultural perspective focuses on the creation of joint communicative and cognitive space and the co-construction of shared knowledge (Hakkarainen, Paavola, Kangas, & Seitamaa-Hakkarainen, 2013). Assessing the shared knowledge construct is difficult in most cases because knowledge structures in an interactive environment are constantly changing. The process of collaboratively constructing a mental model involves the negotiation of different perspectives and the integration of multiple solutions. For example, while participating in collaborative problem-solving tasks, each student in the group may be exposed to various explanations, questions, alternative viewpoints, and evaluations before arriving at a conclusion (Hmelo-Silver & DeSimone, 2013).

The present study adopts the socio-cultural perspective of shared understanding, which regards collaborative group work as a joint learning activity and a process by which a group of children co-constructs a shared mental model. However, there is a scarcity of research in developing reliable measures to assess shared understanding as integrated, coordinated, and joint understanding of a given situation, problem, or task. To track the development of shared understanding in collaborative group work, this study proposes a unique way to evaluate children's causal reasoning, namely the ability to generate *multilink causal reasoning chains*. Multilink causal reasoning demonstrates the ability to produce relational thoughts, to link ideas into frameworks, and to create a detailed flow of reasoning. Successfully determining when to add links to a specific causal reasoning chain as a group constitutes evidence of shared understanding. A reasoning chain that contains links contributed by several students indicates the co-construction of knowledge. The next section introduces multilink causal reasoning.

2.4 Multilink Causal Reasoning

Multilink causal reasoning refers to "the ability to organize incoming information and bridge inferences into coherent causal chains" (Ma, Anderson et al., 2017, p. 71). A functional theory of how a chain of reasoning is created, given by Goodwin and Johnson-Laird (2005), describes a binary relation that consists of two entities and one relation that the entities satisfy. Some relations are transitive, for example, A is better than B and B is better than C. Logically, we can conclude that A is also better than C. Transitive relations are linked together to form a reasoning chain. If we give one entity the role of cause and another entity the role of effect and connect them with a logical relation, then a cause-effect relationship is formed. A sequence of cause-effect relationships forms a causal chain, in which the consequence of each previous link serves as the antecedent of the following link.

The mechanism for the matching of cause and effect is supported by Hummel and Holyoak's (2005) theory of role-based relational reasoning. This theory explains the process of making analogies and creating mental representations based on a computational model called Learning and Inference with Schemas and Analogies (LISA). Hummel, Licato and Bringsjord (2014) extended the LISA model to describe the construction of explanations, which provides a theoretical framework for understanding children's construction of causal chains.

The essential elements of the LISA model include fillers (objects), roles (relations), and role-filler bindings. For example, in the proposition 'people kill wolves', two role-filler bindings can be identified: *kill* (people) and *killed* (wolves). The binding of one relation to one object produces a subproposition in the LISA model (Hummel & Holyoak, 2005). When two subpropositions are both activated in the working memory, a simple relation is formed: *kill* [people, wolf]. The synchronization of two role-filler bindings represents the formation of simple relations. In the extended LISA model that explains cause-and-effect reasoning (Hummel, Licato & Bringsjord, 2014), the filler is not a single object, but a subproposition or a simple relation (i.e., proposition). Imagine that one student said to his classmates regarding the question whether wolves should be killed, "I don't think people should kill wolves because wolf population will go down." This utterance represents one simple relation *kill* (people, wolves) and one subproposition *decrease* (wolf population). The construction of higher-level relations (e.g., causal relation) is expected to follow the same role-filler binding mechanism. However, Hummel, Licato & Bringsjord (2014) believe that a full-fledged synchronization would be very cognitive-resource consuming and perhaps beyond people's working memory capacity. Therefore, they proposed that people categorize propositions or subpropositions into cause or effect groups instead of binding each one with a relational role. Then people only need to connect the cause with the corresponding effect.

If we apply the role-based relational reasoning theory to analyzing children's causal chains, the first step is to define the unit of multilink reasoning, or the link. Each multilink reasoning unit refers to a cause or effect group, which contains a simple relation to be connected with another simple relation via a causal relation. Whether a multilink reasoning unit serves the role of cause or effect depends on how it is connected to another unit, i.e., whether it is an antecedent (cause) or a consequence (effect). So, causal relations explain the between-unit relationships, and simple relations describe the within-unit relationships. Within each multilink reasoning unit, a consequence of the simple relation may also be expressed. For example, students may state that the killing of wolves would lead to a change to the wolf population. This example indicates within-unit causation. Students can either link the simple relations in each unit, or link the consequences to create a chain of reasoning. Between-unit causal relations extend the length of a causal chain. The thoroughness and clarity of the within-unit causal relations determines the potential of extending the reasoning chain. The more thorough the connections within a unit are perceived, the more likely it is that the reasoning chain will be extended.

The complex structure of multilink causal reasoning makes high demands on cognitive processing and requires the ability to discern deep structures despite dissimilarities in surface features (Chi & VanLehn, 2012). The identification of cause-effect relationships relies on considerable amount of practice on connecting ideas and events. Children, especially in underserved communities, often lack experience in connecting ideas, especially when making connections that take several reasoning steps (Flood, 2010). Children also struggle with organizing components within a system and tend to oversimplify complex relations (Assaraf & Orion, 2010; Barman & Mayer, 1994).

In natural social settings, the development of multilink causal chains is highly influenced by the social context in which children receive new information from other people and construct causal mental models through interaction. My previous research has shown that children gain experience with connecting ideas into coherent causal chains by interacting with others. I examined children's production of multilink causal reasoning chains in written policy-decision essays (Ma, Lin, & Anderson, 2013), oral policy-decision interviews (Ma & Anderson, 2015), and oral narratives (Ma, Anderson et al., 2017). Students in these studies participated in a six-week-long unit involving either collaborative group work or direct instruction, focusing on the science and public policy issue of whether a rural community should be permitted to eradicate a pack of wolves.

In the first study, I examined children's ability to generate causal chains in policy decision letters that children wrote to answer the central question of the unit, whether people should be allowed to kill the wolves (Ma, Lin, & Anderson, 2013). I found that the letters of students who participated in collaborative group work contained longer chains of reasoning (many 4-7 link chains) than students who received direct instruction (mostly 1-3 link chains), while most uninstructed control students failed to produce any multilink chains even after controlling for level of conceptual knowledge.

In a second study, I analyzed the production of multilink causal chains in a knowledge transfer task in order to ascertain whether children could produce multilink causal reasoning outside the original learning context (Ma & Anderson, 2015). Students were individually interviewed about whether whaling should be allowed, which is an analogue to the wolf question. Results of this study indicated that students in collaborative groups spontaneously produced more first-order and higher-order concepts and relationships, as well as longer chains of reasoning, than either students who received direct instruction or uninstructed control students, whose performance did not differ.

In a third study, I examined the effect of collaborative group work on children's connection of story events in an oral narrative task, as a far transfer evaluation of children's multilink causal reasoning (Ma, Anderson, et al., 2017). The results of this study indicated that the students who experienced collaborative interaction were more likely to connect separate story events into coherent causal chains, as compared to the stories told by students who received teacher-led direct instruction or who were uninstructed control students.

These studies suggested that children may have developed the ability and disposition to connect ideas after participating in collaborative group work. So, the present study builds on previous work and aims to identify the ways that causal reasoning in the form of multilink causal chains is constructed in Collaborative Reasoning discussions.

2.5 The Rationale for the Present Study

Most educational research relies on end-of-unit or end-of-year tests rather than examining patterns in classroom talk (Cazden, 2001). Surprisingly, even collaborative learning studies may overlook the potential impact of interpersonal communication (Barron, 2003; Ladd, Kochenderfer-Ladd, Visconti, & Ettekal, 2012). In a comprehensive review of literatures on the effects of classroom talk on students' learning in United Kingdom primary schools, Mercer (2008) found that the majority of educational research focuses on the result of classroom talk, rather than the dialogic process in which students engage in step-by-step knowledge construction. However, the temporal dimension of learning cannot be neglected, because "learning is a process that happens over time, and learning is mediated through dialogue, [so] we need to study dialogue over time to understand how learning happens and why certain learning outcomes result" (Mercer, 2008, p. 35). Hence, the **first goal** of this study is to investigate the moment-by-moment internalization of social interaction during collaborative learning. Specially, this study aims to understand the step-by-step construction of multilink causal chains by an individual or by a group of children during Collaborative Reasoning discussions.

Tracking the linking of ideas over the course of a discussion provides a valuable perspective for understanding the temporal development of knowledge construction. However, tracing every idea over an extended period of time presents challenges because the idea may take different forms and the speakers may go off topic at points during the discussion. In a collaborative discussion, speakers frequently interrupt each other and ideas may be fragmentary and scattered. This presents the most formidable challenge in identifying multilink causal chains: any proposition in children's talk may be extended to develop a chain of reasoning but, at the same time, propositions may be stopped from forming into a chain due to interruption, interjection, communication problems, or simply lack of interest. It is almost impossible to identify reasoning patterns in a live discussion and to detect hierarchical cause-effect relationships that are flexible enough to accommodate propositions that address an infinite number of topics.

To overcome this problem, this study uses policy decision essays as the guide to identify patterns of multilink causal reasoning, because they represent the final product of children's thinking and reasoning. The essays were written after students participated in six weeks of collaborative discussions. In the group discussions and the final essays, students were asked to decide whether or not they should allow people in an imaginary town to kill a pack of wolves sighted nearby. Causal chains in the essays are more coherent and better structured than the chains in group discussions. Multilink causal chains take various forms in group discussions due to the constraints of turn-taking. Multilink reasoning chains can be generated by one student or co-constructed by several students in the group. A chain can develop both forward and backward depending on whether the speaker contributes an antecedent or a consequence of the chain currently being developed. Two chains may be intertwined in a single discussion, i.e., one chain may be threaded through the other. Another variation is when a second chain branches off the initial chain. It also frequently occurs that one speaker's causal chain is interrupted by another speaker but is resumed later. For instance, the teacher may prompt further elaboration on a statement, thereby causing the students to return to a previous thought. Although the structure of a reasoning chain can be flexible, the ideas that are used to build the causal chains must be sanctioned by the textbook. In this study, chains that are built entirely on information that comes from outside the textbook are not considered.

The **second goal** of this study is to understand whether group features and individual characteristics affect students' generation of causal chains. A variety of individual-level social and cognitive characteristics are examined in this study, including reading comprehension, basic English proficiency, and several social characteristics such as talkativeness, leadership, having good ideas, and status in peer groups. Previous research has shown that reading comprehension and language proficiency have a big impact on learners' inference generation. Skilled comprehenders were found to generate more explanatory inferences and more coherent situation models in reading comprehension tasks than less skilled comprehenders, and also performed better in their native language than in a second language (Zwaan & Brown, 1996). The reason was that less skilled comprehenders had more difficulty in locating relevant information to make an inference and were less capable of integrating general knowledge and text information than their skilled counterparts (Cain & Oakhill, 1999). Furthermore, language learners' inference generation was often restricted by their level of domain-related knowledge; high-knowledge learners created richer and more accurate mental models than low-knowledge learners (Barry & Lazarte, 1998).

In addition to reading comprehension and English proficiency, social characteristics, including talkativeness, leadership, and status in peer groups, also influence students' academic performance. Previous research has found that popular children are more likely to be perceived as successful students by their peers than socially neglected and rejected children are (Wentzel & Asher, 1995), which contributes to continued academic success. Having reciprocated friendships in school was found to be beneficial to middle school students' social adjustment, which reduced students' emotional distress and improved their academic performance from 6th to 8th grade (Wentzel, Barry, & Caldwell, 2004). Talkative children are often regarded as leaders in class and usually exercise considerable control over the flow of information and enjoy more spontaneous speaking turns during discussion. A study of fourth-graders' leadership development has shown that children who are not quiet and who have good ideas become emergent leaders during collaborative discussions (Li et al., 2007). Emergent leaders are usually more assertive than non-leaders and take more opportunities to speak during the discussion. Talkative children who are at the center of their social group and who enjoy a reputation among their peers for being leaders were also found to take the initiative in incorporating academic words into argumentative talk (Ma, Zhang, et al., 2017).

It is to be expected that socially and cognitively advanced children will play a key role in maintaining successful peer interaction during collaborative discussion by taking the initiative to start a causal chain and helping other students keep up with their thinking. They may provide explanation to less skilled classmates and invite less involved students to speak so that everyone has the chance to share their thinking with the group. They may also resolve conflicts when there are disputed ideas among group members. Socially and cognitively less advanced children are expected to contribute to chain construction by responding to questions or requests raised by socially and cognitively advanced children. The effect of children's social and cognitive characteristics on the formation of reasoning chains is an important component in understanding causal reasoning in collaborative discussions.

The **third goal** of the study is to understand how multilink causal reasoning develops over the course of a discussion. During collaborative discussions, students develop a shared understanding through the maintenance of a "purposeful, shared consciousness" (Mercer, 2008, p. 38) among group members. In this shared communicative space, it is expected that learners work together to evaluate whether specific information is relevant to the main topic, negotiate whether certain links should be included or removed, and discuss in which direction a chain should be extended. As the discussion progresses, conceptual relationships and the words to describe these relationships are expected to become more accessible to students and more widely available for their usage. The shared mental model, therefore, is gradually refined—highly relevant inferences remain in the shared mental model while irrelevant inferences are discarded. The rate of constructing causal chains is expected to increase over the course of discussion as students have more access to the concepts and relationships that are useful in the constructing of reasoning chains.

The fourth goal of the study is to investigate whether patterns of student-student interactions, or occasionally student-teacher interaction, have a proximal effect on students' multilink causal chain construction during collaborative discussions. Chiu and associates (Chiu & Khoo, 2005) invented a new method called *statistical discourse analysis* (SDA) to chart the sequence of events. SDA uses dynamic multilevel analysis to evaluate how previous events affect subsequent events. Chiu (2008) identified four types of social interaction in collaborative problem-solving: supportive actions, critical actions, neutral actions and unresponsive actions. Supportive actions refer to the agreement and acknowledgement of the previous speaker's ideas. Critical actions refer to disagreement and challenges to the previous speaker, neutral actions refer to a neutral reaction or attitude to the previous speaker's position, and unresponsive actions refer to the action of changing the topic. Chiu (2008) investigated whether the interaction pattern between the previous speaker and the current speaker had an impact on the subsequent speaker's production of correct contributions during collaborative problem solving. The results showed that if a speaker politely refuted the previous speaker's statement, it was more likely for the subsequent speaker to produce a correct contribution; however, agreement or rude disagreement during the current turn decreased the likelihood that a correct contribution would be produced in the following speaking turn (Chiu, 2008). Building on Chiu's work, this study employs statistical discourse analysis to investigate the effect of moment-by-moment peer interaction or student-teacher interaction on the construction of multilink causal chains. It is expected that the way a child responds to the previous speaker will affect the likelihood that the next speaking will extend the chain. Agreement between speakers is expected to extend the chain of reasoning, while disagreement between speakers is likely to arrest the development of the chain.

This study also identifies leadership moves which enable teachers and students to manage discussion and to facilitate the development of reasoned argumentation. Previous studies have documented emergent leadership in collaborative learning groups and found widely distributed leadership functions among group members (Gressick & Derry, 2010; Li et al., 2007; Mercier, Higgins, & Da Costa, 2014; Sun, Anderson, Perry, & Lin, 2017). However, these studies focus on the overall effectiveness of leadership moves but not the proximal effect of leadership moves on children's thinking and reasoning during collaborative discourse. The present study uses statistical discourse analysis to analyze the proximal effect of different types of teachers' and students' leadership moves on a child's construction of a causal chain. It is expected that talkative and socially central children will invite quiet and socially peripheral children to participate, thereby increasing the likelihood that they will contribute links to a causal chain. When a teacher intervenes and prompts a student to express his or her opinion, and to use reasons or evidence, it is likely that the student will extend the chain of reasoning.

The fifth goal of the study is to investigate whether a student's ability to initiate or add links to causal chains during collaborative discussion increases his or her likelihood to generate causal chains in the final policy decision essay. The spread of ideas from the group discussion to the individual essay indicates knowledge transfer. According to Day and Goldstone (2012), transfer is the "recruitment of previously known, structured symbolic representations in the service of understanding and making inferences about new, structurally similar cases" (p. 154).

Prior knowledge and structural similarity both have a significant impact on whether learners can transfer between contexts. Learners interpret a new problem by drawing on prior knowledge and they understand how the new problem is related to what they already know, which they reinterpret and apply to the new problem. Different people create different conceptualizations of the same problem, and in turn take different routes to solve the problem. What is to be retrieved is highly dependent on how a learner understands the problem and the way he or she prioritizes information. The prior knowledge that a novice learner retrieves often has a lower cueing priority than the knowledge retrieved by an expert learner, but the knowledge appears to be relevant and may be the most accessible knowledge to the novice learner (Chi, 2006). There are four factors that influence knowledge transfer: prior knowledge, structural relations, concreteness of description, and the learners' cognitive load during information processing (Day & Goldstone, 2012).

First, prior knowledge plays an important role in transfer because it determines what kind of information will be retrieved and later applied to the new situation (Bransford & Schwartz, 1999). Abundant exposure to a specific knowledge domain is beneficial to learners for developing a general understanding of the field and forming a coherent mental model. Second, the explicitness of structural relations affects the readiness of transfer. Connectives, relational markers, and explicit labeling of relations may help learners to identify the structural relations more easily and process the mapping more quickly. Third, the concreteness and comprehensibility of information acts as a facilitator to transfer. Learners, especially novices, may spend more time in decoding information and may encounter processing difficulty when there are unnecessary contextual details or confusing information in the text. Fourth, the influence of cognitive load cannot be neglected because processing structural relations requires considerable effort and attention. Learners have to complete surface level processing before processing higher-level information.

As a result, transfer is more likely to happen when learners are exposed to a similar relational structure to the ones with which they are already familiar or have some generalizable knowledge. The transfer of multilink causal chains requires sufficient background knowledge and a general picture of how concepts are related. Students with abundant prior knowledge and those who have mastered the causal structure are likely to transfer the mental model to a similar context. Therefore, it is expected that children who generate multilink causal chains, especially those who initiated causal chains during discussion, will produce causal chains in their individual essays. Students who did not produce any causal chains during the discussion, but have been exposed to certain types of causal structures when listening to peers, are also expected to produce causal chains in the individual essay. However, those who did not produce any causal chains during the discussion, nor had the chance to sit with those children who produced causal chains, are expected to have difficulty in producing causal chains in their essay.

To summarize, the present study analyzes the moment-by-moment construction of causal chains and aims to identify the recurrent patterns of productive social interaction that enhance children's ability and disposition to connect ideas. Specifically, this study attempts to answer five research questions:

1. what types of multilink causal chains are produced in collaborative discussions?
2. what group features and individual characteristics increase the likelihood of students generating causal chains?
3. how does the construction of multilink causal chains develop over time?
4. how does student-student interaction and occasional student-teacher interaction affect the construction of multilink causal chains?
5. does a student's ability to initiate or add links to causal chains in group discussion increase his or her likelihood that he or she will produce causal chains in the individual essay that follows the group discussion?

CHAPTER 3

METHOD

3.1 Design and Procedure

In this study, students from two underserved populations—African American and Hispanic American—participated in what was called the Mindful Instruction Project, which aims at improving conceptual understanding, motivation, and engagement, as well as the language, thinking, and reasoning of underserved children. The Mindful Instruction Project was implemented in two waves in the course of two academic years, each wave being a cross-section of the whole study. Children were immersed in a specially-developed six-week Wolf Reintroduction and Management Unit (hereafter referred to as Wolf Unit) that addressed a socio-scientific policy issue in an imagined community, where local people are concerned about potential threats of a pack of wolves and must decide whether they should hire professional hunters to kill the wolves (Jadallah et al., 2009). Three intertwined subsystems of knowledge were introduced in this unit—ecosystem, economy, and public policy. The text was written in various genres, and included expository texts, newspaper articles, and formal letters. These texts presented both sides of the controversial issue and explicitly asked the students to weigh and evaluate each side. For example, students could conclude that the wolves' presence in town attracts tourists and boosts the local economy, but also poses a threat to the safety of campers and other visitors.

Classrooms were placed in two treatment conditions and taught the material from the Wolf Unit via two contrasting instructional approaches: collaborative group work and whole-class direct instruction. A wait-listed control condition was set up to provide a baseline to gauge the effects of the two treatments; this condition consisted of classrooms whose learning of the

Wolf Unit material was deferred until the entire intervention was concluded. Due to the fact that students who received direct instruction had limited opportunities to speak, and students in the control group were videotaped in only one class session prior to the Wolf Unit, this study focuses on the data collected from the students who learned the material via collaborative group work.

The students in Collaborative Group work (CG) classrooms were arranged in heterogeneous groups (called expert groups), each of which studied a subsystem of knowledge through peer-managed collaborative reading, discussion and other activities such as graphing, calculation and information research. At the beginning of the intervention, each expert group had an initial Collaborative Reasoning discussion of the 'big question': whether local people should hire professional hunters to kill the wolves. In order to answer this question, each expert group focused their efforts on developing expertise on the ecosystem, economy, or public policy. At the conclusion of the Wolf Unit, each expert group made a poster presentation to share what they had learned with the rest of the class. Finally, students reconsidered the 'big question' in a second Collaborative Reasoning discussion. For this final discussion, groups consisted of a combination of experts from all three subsystems. This jigsaw structure (Aronson & Patnoe, 1997) enabled students to discuss the 'big question' with peers who had studied the problem from within a different knowledge domain.

3.2 Participants

The collaborative group work condition was comprised of 257 students who were recruited from 12 classrooms in eight public elementary schools in northern and central Illinois. A group of students, called target students, was videotaped every day during the intervention. The target students formed a cross-section of the classroom in terms of gender and ethnicity, as well as talkativeness and academic ability, based on teacher judgment of students' social and cognitive attributes. Non-target students were only videotaped during the two Collaborative Reasoning (CR) discussions—the initial discussion that took place before students studied the Wolf Unit material and the second one that concluded the Unit.

A preliminary examination of the video clips revealed that the amount of student talk varied from day to day depending on the topic and task of that day. For example, on some days, students spent the majority of the time reading texts rather than discussing the ‘big question’ or any closely related topics. Similarly, much of the student communication during the poster-making sessions consisted of task-fulfilling utterances (e.g. “hand me the scissors,” “put that one here”) rather than discussion of the ‘big question.’ Not surprisingly, the two CR discussions produced a significant amount of talk directed toward the ‘big question’ and involved a more comprehensive consideration of the concepts and information from the unit. The final CR discussion, in particular, contains the highest density of academic vocabulary, and provides the most valuable data for the analysis of students’ cognitive development.

I conducted an academic vocabulary search (the list was generated based on textbook glossary) in the transcripts of 72 four-minute videos that were systematically sampled from six Wolf Unit sessions. In chronological order the sessions from which the videos were taken are: the introduction to the Wolf Unit, the initial CR discussion, the session on wolves in the United States, a lesson on the ecosystem, the poster-making session, the session in which students presented their posters using their newly acquired knowledge domain to the entire class, and the final CR discussion. The average use of academic vocabulary in each session, which is one indicator of students’ content talk, is presented in Figure 1. It shows that students’ content talk gradually increased as they acquired more wolf-related knowledge, with the most content talk occurring in the second CR discussion.

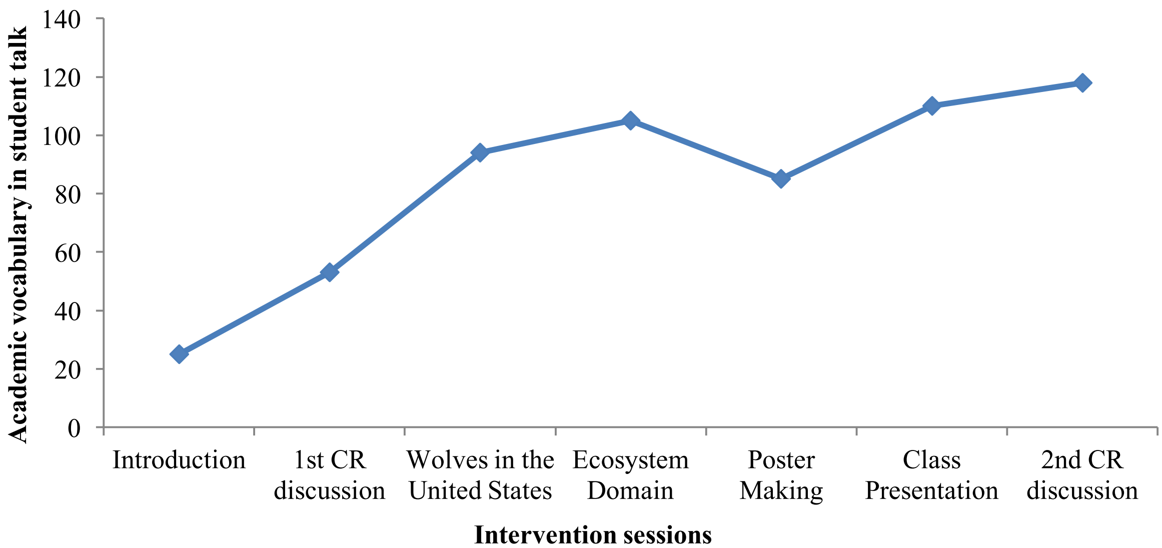

Figure 1. Use of academic vocabulary over six intervention sessions.

In the final Collaborative Reasoning discussions, two or three children from each of the three expert groups (ecosystem, economy, and public policy) were selected to form a new group to discuss whether people should be allowed to hire professional hunters to kill the wolves. This yielded 39 discussions (3-4 group discussions from each classroom). However, due to the usability of the video data (e.g., not all parents gave consent to use video data of their child) and time limitations of the project, a total of 24 group discussions were selected for analysis (two groups from each classroom). A total of 162 students participated in these 24 collaborative discussions that provided 389 minutes of talking. Two of the 162 students were not included in the data analysis due to lack of parental consent. Therefore, the final dataset contains 160 students, 93 girls and 67 boys. Most of the students are African American (44%) and Hispanic American (46%), with the rest being European American (8%) or biracial (2%). Additionally, eleven of the students (7%) were enrolled in Individualized Education Program (IEP) to receive special services such as reading lesson or speech therapy at the time of the intervention. Sixty students were in the Ecosystem expert groups, 50 students were in the Economy expert groups, and 50 students were from the Public Policy expert groups. Table 1 shows the number of students in each group and the duration of the group's discussion session.

Table 1

Number of Children and Duration of Talk in Collaborative Groups

Year	**Class**	**Dominant ethnicity**	**Number of students**		**Duration of talk (min: sec)**	
			Group 1	Group 2	Group 1	Group 2
Y1	Class 1	Latino/a	7	6	15:35	25:01
Y1	Class 2	Latino/a	7	6	17:09	14:15
Y1	Class 3	Latino/a	7	6	18:32	14:16
Y1	Class 4	African American	5	4	21:00	19:34
Y1	Class 5	African American	6	10	14:31	14:27
Y1	Class 6	African American	5	5	20:36	18.39
Y2	Class 7	Latino/a	7	6	16:36	14:04
Y2	Class 8	Latino/a	7	7	12:14	18:21
Y2	Class 9	Latino/a	9	9	14:58	12:24
Y2	Class 10	African American	8	9	25:45	12:24
Y2	Class 11	African American	6	6	15:00	15:43
Y2	Class 12	African American	4	8	10:39	07:22
Total		12	160		389	

3.3 Discussion Transcripts

I transcribed the 24 final CR discussions in their entirety. The transcription was reviewed for accuracy by two native speakers of English. One reviewer is African American and the other reviewer is Hispanic American—both have expertise in capturing the linguistic features of the dialect. We discussed the discrepancies in the transcription and collaboratively produced a corrected version. The transcription was completed using InqScribe Digital Media Transcription Software to record speech verbatim, overlapping speech, intervals within and between speaking turns (e.g., pauses, interruptions, or contiguous utterances), significant nonverbal behavior (e.g., raising hands before speaking), timestamps in the beginning and at the end of each speaking turn, and transcriber comments. A basic communication unit is a speaking turn, which consists of a set of propositions that one speaker delivers while holding the floor. Speaking turns are further categorized into four types: full turns, simultaneous turns, interjections, and interruptions (Chinn et al., 2001).

Full turn. A full turn is when the speaker holds the floor and expresses at least one complete idea. The situation in which a speaker is credited a full turn includes those instances when [1] the speaker holds the entire floor and expresses one or more complete ideas without being interrupted by other speakers; [2] the speaker begins the turn and holds the floor for a short period before being interrupted, but is able to finish his/her idea without being prevented by the interrupter. In situation [2], if the *second* speaker (the interrupter) is able to express one or more ideas after interrupting the first speaker, the turn is coded as a full turn for the second speaker. However, if the *second* speaker fails to gain the floor after interrupting the first speaker, this is coded as an interjection. As a note, polling turns, in which students successively give their response (agree or disagree) to the central question, were coded as *simultaneous turns* by Lin et al (2012); however, in this study, polling turns are coded as *full turns*.

In the following example, Sarah held the floor and finished one complete sentence before Maya jumped in, but Maya only added one word "buses" in order to suggest that children can take the bus to avoid being attacked by wolves. Sarah continued her speaking turn, during which Maya interpolated one word "community" to contribute to Sarah's statement. Towards the end of Sarah's turn, both Demisha and Maya contributed to Sarah's idea without taking over the floor. Therefore, Sarah's turn was coded as a *full turn* while Maya and Demisha's utterances were coded as *interjections*. In the following excerpt, each speech act is indicated by [Number] in order to chart the patterns of overlapping episodes or interjections. Lowercase letters refer to different endpoints in multiple turns.

Sarah: [chuckles] I think (?) the best decision for for the the town for the town is like you know the schools when um parents are gettin' scared for their children to to walk to school now because they're scared that a wolf'll come after 'em. [1] [1] I think they should like move like move a school to move a school really close close to the town like [2] [2] Yeah, the community where everybody lives so they don't have to worry. And like have like somebody out there watchin' all the kids walking in the [3] [4] [5] [4] [3] (???) [5a] yeah. [5b]

Maya: [1] Buses. [1]
Maya: [2] Community. [2]
Demisha: [3] Like uh [4] [5] like uh [4] City Block Program. [3] [5b]
Maya: [4] Like [5] security guard. [4] [5a]

Simultaneous turn. A *simultaneous turn* is identified when two or more speakers express complete ideas at the same time. The most common occurrence is that several students enact one full turn simultaneously in response to a question or an expressed opinion (this was coded as a *group turn* in Chinn et al., 2001). In this case, the speakers are usually unidentifiable (transcriber uses *students* instead of specific names to record the speakers). Another type of *simultaneous turn* is when two speakers simultaneously interrupt the student who is holding the floor and each new speaker immediately and simultaneously completes a full turn. Here is an example of a simultaneous turn.

Demisha: Because they said there's no actual record of a healthy wolf [1] trying to kill a human [1] human being. So if they weren't actually trying to kill a human being, there's no real harm to humans so why would you want to have hunters come?

Students: [1] Trying to kill a human being. [1] *simultaneous chain made by several students*

Interjections. Interjections are turns in which the speaker fails to gain or hold the floor. As defined by Chinn et al (2001), interjections include [1] fragmentary turns (back channeling or interjected comments); and [2] incomplete turns (failed attempts to gain the floor with or without interruption). An incomplete turn occurs when a speaker does not form a complete thought before beginning to speak and consequently has to give up the floor to rethink his/her statement, or more frequently, when the speaker is interrupted by a different speaker, which leads to a new speaking turn. In the following example, Sarah was interrupted by Maya (the interruption is marked as //); however, Sarah gained the floor because Maya forgot what she was going to say.

Sarah: I I th- *failed attempts to gain the floor with interruption*
Maya: I think the reason is – Go ahead.
failed attempts to gain the floor w/o interruption
Sarah: That's OK.
Maya: No. I just forget what I'm going to say.

Sarah: And maybe build a like a home or program for them to study them [1] [1] and what they need and like do they really need the trees so they can study more about them.
Maya: [1] You mean like um a nature preserve? [1] *interjection comments*
Onna: Or like a- [2] Or like a- [2] *incomplete turn*
Edward: [2] Yeah. A nature preserve. [2]
Maya: That would actually [3] work [3] really well.
Sarah: [3] Yeah. [3] *back channeling*

Interruptions. An interruption is when a speaker stops the continuous progress of another speaker's turn. The results of interruptions include [1] the speaking turn immediately being taken over by the interrupter; [2] the floor being regained by the original speaker after being interrupted; [3] the speaker being able to finish the sentence despite the turn being taken over by the interrupter. In small-group discussions, ideas are often distributed over more than one speaking turn. If the second speaker interrupts the first speaker by bringing up a new idea, by managing the flow of discussion, or by going off topic, the floor is usually taken over by the second speaker. However, if the second speaker asks a clarification question or continues the first speaker's idea by supporting or challenging it, the first speaker often retains the floor.

Descriptive statistics of discussion corpus. The whole corpus is 389 minutes long and includes 4,013 turns of speaking, among which 3,354 turns contain at least one complete idea, and 659 turns are interruptions, interjections, back channeling, or incomplete turns. Approximately 94% of the talk (3,161 speaking turns) is contributed by the students and the rest is from teachers or participant observers who videotaped the discussion (193 turns of speaking). The distribution of speaking turns in each discussion group is presented in Table 2.

Table 2

Distribution of Speaking Turns in Each Discussion Group

Group	Total Number of turns	Full turns	Interjections	Interruptions	Student full turns	Teacher full turns
First wave of the intervention						
1	123	88	34	1	84	4
2	221	183	30	8	175	8
3	124	117	4	3	116	1
4	141	110	28	3	109	1
5	252	204	43	5	197	7
6	376	334	40	2	315	19
7	108	94	11	3	94	0
8	119	101	17	1	101	0
9	116	84	30	2	68	16
10	100	81	19	0	76	5
11	136	98	36	2	96	2
12	147	137	9	1	121	16
Second wave of the intervention						
13	273	214	57	2	212	2
14	175	159	16	0	158	1
15	301	234	64	3	227	7
16	126	91	33	2	76	15
17	57	49	7	1	45	4
18	44	40	4	0	34	6
19	63	58	5	0	41	17
20	97	89	8	0	76	13
21	186	157	27	2	145	12
22	236	194	41	1	174	20
23	222	186	33	3	178	8
24	270	252	18	0	243	9
Total	**4013**	**3354**	**614**	**45**	**3161**	**193**

3.4 Assessments

3.4.1. Pretest measures

Pretests assessed reading comprehension and oral language proficiency. The Gates-MacGinitie reading comprehension test (fourth edition, level five, form T; MacGinitie,

MacGinitie, Maria, & Dreyer, 2000) was used to assess students' level of independent reading ability. The test contained 48 multiple-choice questions following short passages. Students were given 35 minutes to complete as many questions as they can. Raw reading scores included the number of test items that a student attempted to answer, the number of correctly answered items, and the number of incorrectly answered items. The final scores were derived by reducing the correct scores due to guessing (Right - Wrong/3).

The rapid automatized naming task (Snodgrass & Vanderwart, 1980) was used to assess student's oral English fluency. Students were pulled out into the hallway to look through two pages, each of which contained 24 pictures of simple objects like bus, rabbit, and clock. They were asked to name these objects as fast as possible and were allowed to skip pictures if they do not know the object. Examiners audiotaped the whole naming process, recorded the time that each student used to name two sets of pictures, and made notes on events such as skipping pictures, long pausing, or giving a different name for the picture. The final score was the number of words that students correctly named per minute.

Students filled out a social questionnaire before the intervention to indicate which peers they considered to have the most things to say during class discussions, to be too quiet, to be good leaders, to have good ideas, and which students they liked to play with at school (friendship). Talkativeness was calculated by deducting the number of classmates' nominations for being too quiet from the number of nominations for usually having the most things to say during classroom discussions, then divided by the total number of students in the classroom. Additionally, students rated how much they liked to play with each classmate, using a five-point Likert scale. Three centrality scores were derived from friendship nominations using social network analysis (Butts, 2008) through UCINET software (Borgatti, Everett, & Freeman, 2002);

the scores were degree centrality, betweenness centrality, and information centrality. Degree centrality represents the number of connections one individual has with others in the network, which includes both in-degree and out-degree nominations. Betweenness centrality indicates how many unconnected pairs of people can reach one another in the minimal number of steps by going through another individual (Wasserman & Faust, 1994). Information centrality is calculated as the harmonic mean of the information of the paths between a given point and all other points to which it is directly or indirectly connected in a social network (Stephenson & Zelen, 1989). Information of a path is the reciprocal of the distance between two points. A point with high information centrality is more important and enjoys higher status in the social network than points with low information centrality. All these social measures are normalized by class mean in order to enable between-class comparisons.

3.4.2. Posttest measure

After the intervention, students were given 40 minutes to write a policy decision letter regarding the 'big question': *do you permit the Winona County Board to hire a professional hunter to kill the pack of wolves?* The examiner explained the writing activity to the whole class and gave students 10 minutes to write an outline. Then the examiner distributed a response letter template to students and they began writing. Among the 160 students who participated in collaborative discussion, 154 students wrote the decision letter. The six students who did not write the letter were absent during the day when the written test was administered, including four girls and two boys. Five of them were African American and the other student was bi-racial. Three of the absent student were from the same class while the other three students were from three different classes, among which two students enrolled in Individualized Education Program.

3.5 Coding Multilink Causal Reasoning in Collaborative Discussions

3.5.1. Identifying patterns of multilink causal reasoning

Based on a bottom-up search of students' decision letters, seven multilink reasoning models were identified. These are the *ecosystem-economy model,* the *ecosystem-oxygen model,* the *leftover model,* the *compensation model*, the *supply model*, the *competition model* and the *tourism model.* I searched the children's group discussions to identify similar patterns of multilink reasoning models as those that appear in the wolf decision letters. Chains that do not follow the exact patterns of the seven reasoning models above are categorized as *novel chains*. While still holding a causal connection with the first several links of one of the reasoning models, the final links in *novel chains* diverge from the trajectory of these seven reasoning models. Chains that are built entirely on information that comes from outside the Wolf Unit are not considered. Each model is explained in detail in the following sections.

The *ecosystem-economy model* contains eight links, which make it the most complicated reasoning chain. Figure 2 illustrates the arrangement of the causal links in this model. A multilink causal reasoning chain is considered to be complete when a possible consequence of a simple relation is verbalized in each link. For example, in the ecosystem-economy model the first link contained the simple relation: "People kill wolves." The consequence of this simple relation was expressed as: "People killing wolves will cause a change to the wolf population." In the second link of this ecosystem-economy model, the simple relation was "Wolves eat elk", and the consequence was expressed as: "The elk population will change." These two links are connected if the two expressed consequences are considered to form their own causal relation. Causal links formed in the same way extend the reasoning chain, in the case of the ecosystem-economy model, up to eight links.

The ecosystem-economy model includes both concrete and abstract concepts, lower-order and higher-order relationships, as well as schematic knowledge such as the concept of the food chain and the interdependence of businesses. As shown in Figure 2, the chain is divided into eight links (A1→A2→A3→A4→A5→A6→A7→A8). Within-link causation occurs when a simple relation is connected with a consequence. Between-link causation occurs when a consequence is connected with a consequence from another link. Between-link causal relations extend the length of a causal chain. The thoroughness and clarity of the within-link causal relations determine the potential for extending the reasoning chain. The more clarity and thoroughness with which the connections within a unit are perceived, the more likely it is that the reasoning chain will be extended.

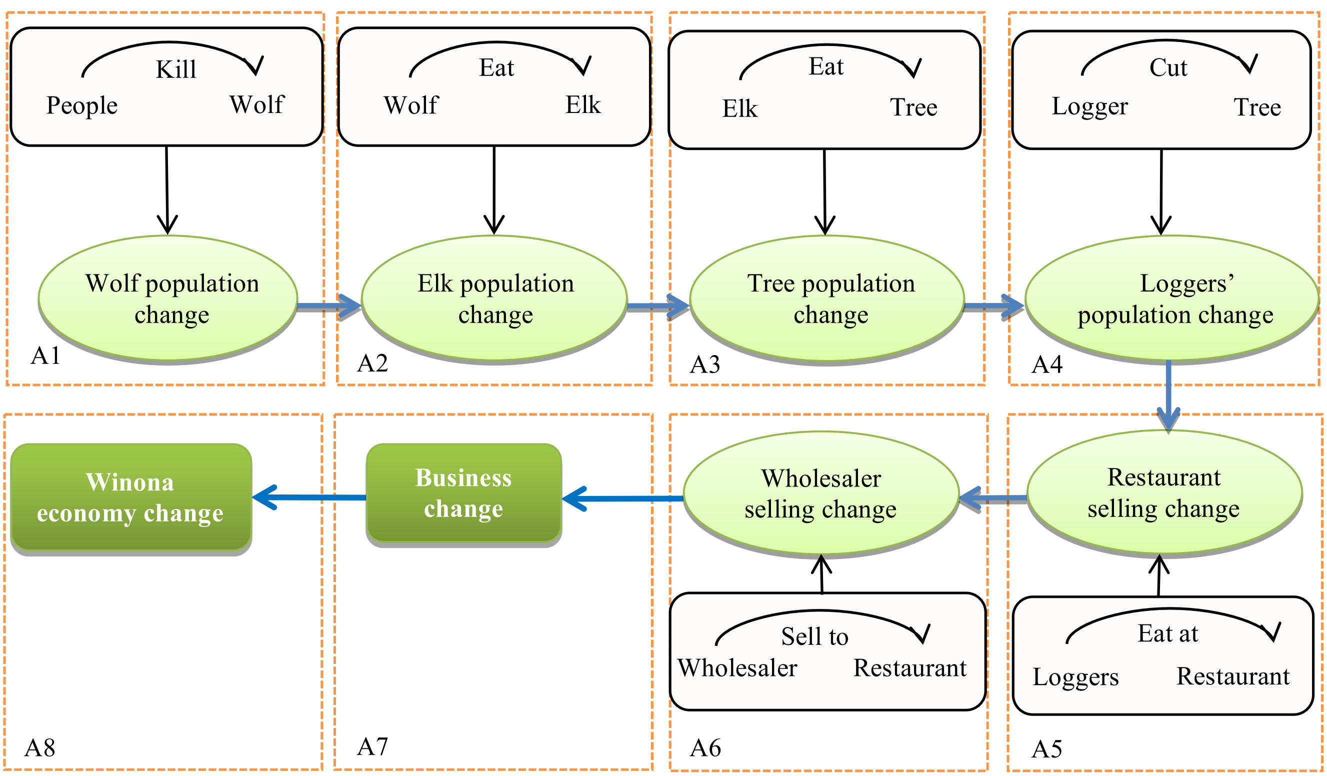

Figure 2. The multilink reasoning structure of ecosystem-economy model (Model A).

The *ecosystem-oxygen model* contains five reasoning links. As shown in Figure 3, the first three links are the same as in the ecosystem-economy model; however, the ecosystem-oxygen model remains within the ecosystem domain without spreading into the economy domain. Students stated that the decrease of tree population will cause “less oxygen” on earth because “trees produce oxygen.” The next consequence in this chain was “if there is less oxygen for people to breathe, people won’t be able to stay alive.” Although this chain is flawed by overgeneralization, it nevertheless demonstrates causal reasoning and shows children’s disposition toward finding relationships between concepts. It must be understood that the focus of this study is not to ascertain how accurately a child can recall information, but rather how well they can connect ideas into a chain of reasoning.

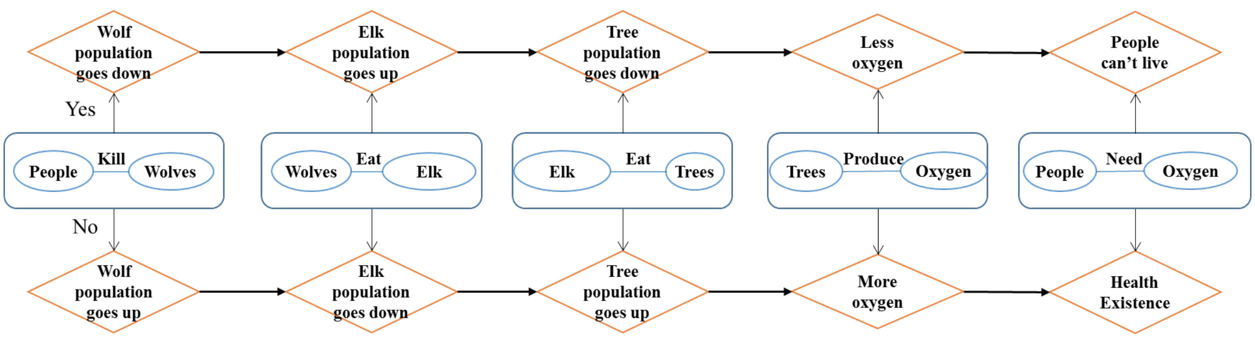

Figure 3. The multilink reasoning structure of the ecosystem-oxygen model (Model B).

The *leftover model* also falls within the ecosystem domain. It contains four reasoning links (see Figure 4). The first link contains a within-link causation: “If people kill wolves, the wolf population will go down.” Students learned from the Wolf Unit texts that wolves eat other animals and often leave leftovers for scavengers. This information produced the consequence: “If people kill wolves, there will be less food for other consumers who rely on wolves’ leftovers to survive.” The following consequence was: “These animals might die from starvation.” As can be seen, the leftover model describes the basic idea of the food chain.

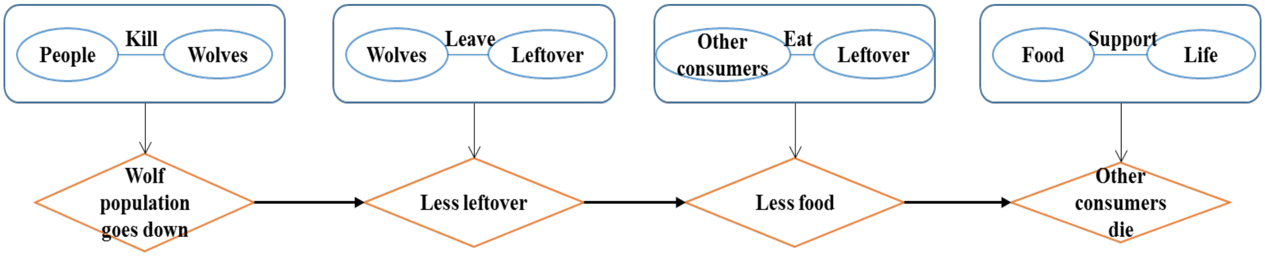

Figure 4. The multilink reasoning structure of the leftover model (Model C).

The *competition model* is another chain that spreads over the ecosystem and economy domains. The main idea of this reasoning chain is that hunters and wolves are competitors (see Figure 5). If the animals that professional hunters target are wolves' prey, an increase in the wolf population will negatively affect hunters' profits. As a result, hunters may be unable to support their family. As one student stated, "You know, if the wolves keep killing the elks, and every animal, the hunter can't go hunting no more and it's affecting their lives and then affecting everybody else's."

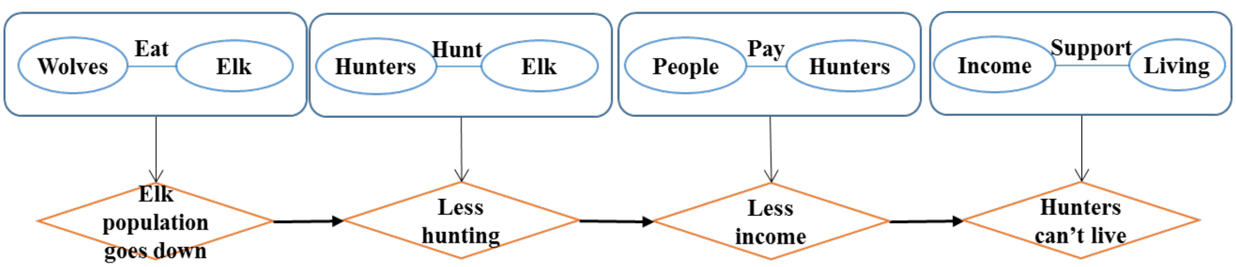

Figure 5. The multilink reasoning structure of competition model (Model D).

The *compensation model* falls within the economy domain and contains four reasoning links, as presented in Figure 6. This chain starts with the fact that wolves kill livestock, and proceeds to the consequence that, as a result, ranchers lose money. Some students stopped at this link. Other students progressed further and concluded that ranchers would not lose money if they could prove their loss. The consequence was stated in this way: "Ranchers need to show proof of losing livestock to be compensated by the government."

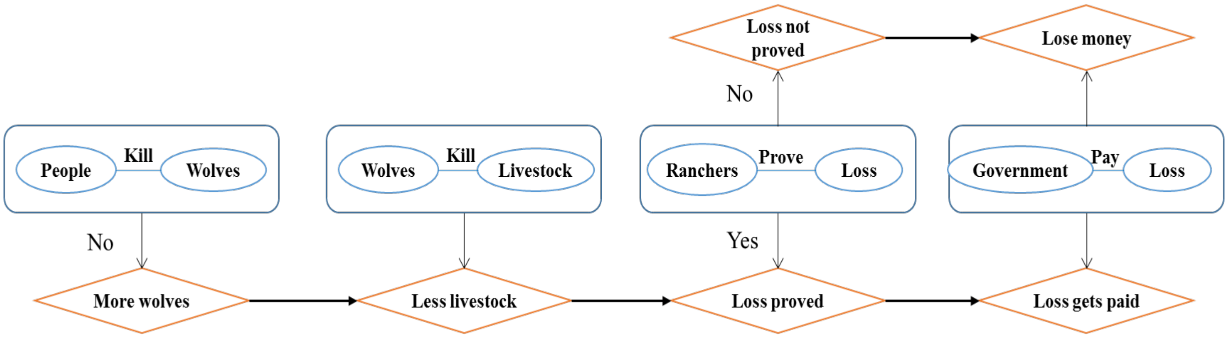

Figure 6. The multilink reasoning structure of the compensation model (Model E).

The *supply model* falls within the economy domain and starts with the fact that wolves eat livestock (see Figure 7); however, it proceeds in a different direction than the compensation model. It contains the simple relation that restaurants buy meat from ranchers, and therefore a result of ranchers losing livestock would be a lack of meat for restaurants. The subsequent consequence was that restaurants' business would be negatively affected by the shortage in supply. This chain contains five reasoning links.

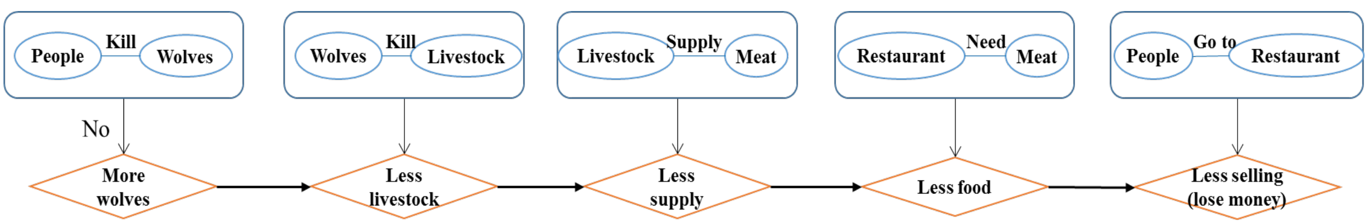

Figure 7. The multilink reasoning structure of the supply model (Model F).

The *tourism model* is another chain in the economy domain and contains three reasoning links. The tourism model has two types. The first type (see branch 1 in Figure 8) builds on the idea that wolves contribute to local tourism by attracting visitors interested in seeing them, as one student stated: "As there are wolves in the town, tourists will come to see wolves and spend money on food or supplies; if people kill wolves, fewer tourists will visit, which will hurt tourism." In contrast, some students argued that wolves would hurt tourism because people would be scared by them and would stop coming. This thinking produces the second type of tourism model (see branch 2 in Figure 8).

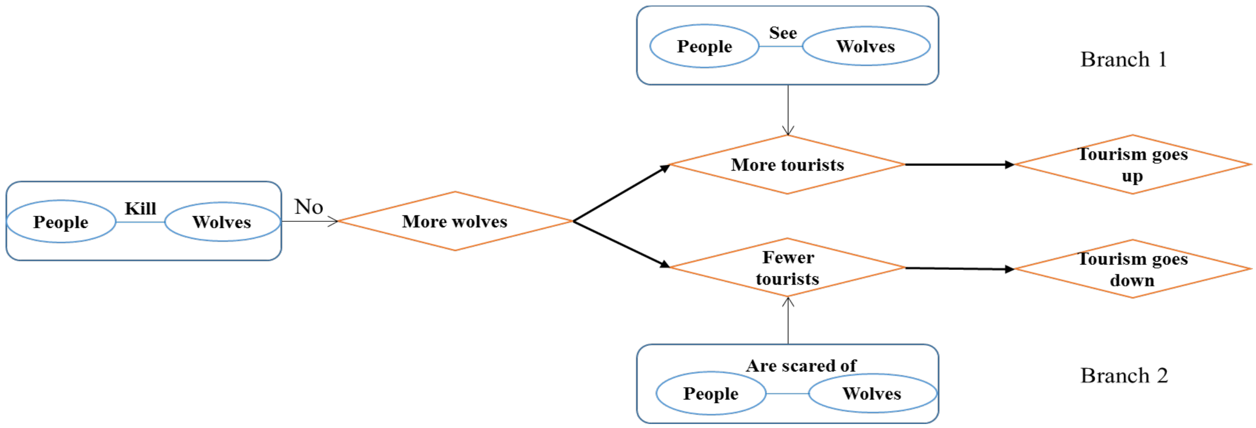

Figure 8. The multilink reasoning structure of the tourism model (Model G).

Seven types of multilink causal reasoning models and novel chains were identified in the 24 discussion transcripts. Since some speakers produced more than one causal chain within a speaking turn, and some causal chains spread over multiple speaking turns and across multiple speakers, a binary code was assigned to each speaking turn to indicate whether the turn contained one or more links belonging to a single chain. The coder further identified the number of links and the type of chains. I coded all the transcripts and a second coder independently coded 20% of the transcripts to ensure the reliability of the coding scheme. The intercoder reliability indicated high agreement between coders (percentage of agreement was 95% and Cohen's K = .71).

3.5.2. Identifying structures of multilink causal chains

The structure of multilink causal chains varies according to the arrangement of turn-taking during collaborative discussions. The following section summarizes the different layouts of the multilink causal chains that occurred in the group discussions. In the following description, the multilink reasoning models are denoted by uppercase letters indicating the order of appearance (A-G). Each reasoning link is numbered. For example, B1 refers to the first link in the *ecosystem-oxygen model*. Cause and effect are identified in each link if the child explicitly describes the simple relation and its consequence. For example, **C2-cause** indicates that the utterance functions as a cause; **C** indicates the *leftover model* and **2** indicates the second link.

Straight chain. A straight chain is a series of consecutive and interlocked causal links. In the following excerpt, Sophie gives her position on the 'big question', which is that people should not kill wolves. Prompted by Eloy, Sophie further provides a reason for her argument and creates a within-unit causal relation. Then another student, Justine, constructs a five-link straight chain and produces what we have already referred to as the ecosystem-oxygen model. There are four between-link causal relations and one within-link causal relation in Justine's chain.

Sophie: I think that we shouldn't [kill wolves] [**B1-cause**].
Eloy: Why?
Sophie: Because you just hurting the, population of of the wolves [**B2-effect**].
Justine: I say they shouldn't [**B1**] because, um, the wolves kill the elk [**B2-cause**], and when the elks are dead, [**B2-effect**] there are more trees, [**B3-effect**] so there's more oxygen for us. [**B4-effect**]
Monique: I was going to say that.
Justine: So the wolves make it better for us [**B5-effect**].

Keil (2006) is among those who have noticed that people's conceptualization of causal structure is often implicit rather than expressed, especially in conversations. In the above example, Justine implies two unstated but presumed relations: *elk eat trees* and *trees produce oxygen*. In the discussion, nobody challenged the speaker on these two relations or asked for clarification, which indicates that there is implicit agreement between the speaker and the audience. Chains that contain implicit links are coded in the same way as chains that contain explicit links if it is evident that group members have a shared understanding of the underlying relations.

Branched chain. A branched chain originates from the middle of another chain and produces a different conclusion. The branch is most often initiated by another speaker, though it is also possible for one speaker to construct a branched chain by identifying two or more effects of a given cause. As shown in the following example, Tevon initiates a chain by stating his position on the 'big question'. The chain is interrupted by Raven and Enrique after the first link ("[hunters] should not just get wolves"), but is resumed later when the teacher asks the group to stay on topic. Tevon proceeds to develop the chain when he gains the floor. Enrique agrees with Tevon's statements and immediately reiterates Tevon's chain, but also initiates a branched chain within the same tourism model when he concludes that tourists may be scared by wolves and as a result decide not to go camping (marked by number 7). The branched chain is graphed in this way: G3b←G2b←G1→G2a→G3a.

Tevon: I think they said only if um ... I forgot, well, I think they should not just get wolves [**G1**]. They own space.
Raven: I think they should kill 'em because then people ain't gonna be able ta go out to hunt and like they want to [interrupted by Enrique].
Enrique: Wel- they probably ain't got they huntin' license. They can go to jail for that. Not having a hunting license?
Teacher: Topic, on topic.
Tevon: Um like on, our poster it said that it could help tourism go up [**G3a**] because people will come- wanna come see the wolves [**G2a**].
Enrique: Uh our book's on tourism. <u>No one will, go camping [**G3b**]</u>, <u>they will be scared [**G2b**]</u>, and then on the second one we put might go up [**G3a**]. People want to see the wolves too at the same time [**G2a**] so maybe. ... 'Cuz people wanna see the wolves [**G2a**], but, some people might wanna see the wolves [**G2a**], <u>some people might be scared [**G2b**]</u>.

Threaded chains. The threading phenomenon is often observed in non-real time interaction such as online discussion or messaging (Black, Levin, Mehan, & Quinn, 1983). However, threads of conversations are also likely to appear in group discussions with at least four speakers because at least two subgroups can be formed and subgroups may pursue different topics simultaneously. The simple version of threading is when each subgroup focuses on one line of reasoning without engaging other lines but intertwines their own line with other subgroups' speaking turns. The complex version of threading is when speakers contribute to multiple topics, making it difficult to differentiate one thread from another. In threaded chains, links are interrupted by multiple speakers at different points but the links are eventually reunited to form a complete chain. In the discussion transcripts, chains were not threaded by other chains, but were threaded by speaking exchanges that do not contribute to any chain of reasoning. The following excerpt provides an example. Olivia generates the first three links in the ecosystem-oxygen model, and Akira adds the fifth link to this chain, but only presumes the fourth link "trees produce oxygen" without explicitly stating it. Rafael jumps in and breaks the chain by inviting Emilio to contribute some thoughts. Olivia does not give the floor to Emilio but adds a new link to his original ecosystem-oxygen model. The new link branches off the ecosystem-oxygen model built by Akira and Olivia and also begins a new chain. In the next turn, Emilio takes the floor to respond to Rafael's invitation from two turns ago. Rafael and Emilio's turns thread into the chain that is developed by Olivia and Akira.

> **Olivia**: Cause if the wolves are gone, if the wolves are gone [**B1-effect**], the elk is gonna keep on eating the um um plants [**B3-cause**] because there is nobody to kill them [**B2-cause**] and then we're not gonna have enough [**B3-effect**] because the elk is eating the plants [**B3-cause**]. Not enough plants [**B3-effect**]. (Thread 1)
>
> **Akira**: We need oxygen [**B5**]. They have plants. They have food. (Thread 1)
>
> **Rafael**: Have you thought something yet? Emilio, have you thought something yet? (Thread 2)
>
> **Olivia**: We do eat plants [**B4-new link**]. We are not going anybody. We eat stuff that grow from trees [**B4-new link**]. (Thread 1)
>
> **Emilio**: Um I think no [Rafael: Shhh-Shhh-Shhh-Shhh] 'cause they 'cause they gotta eat just like we gotta eat. (Thread 2)
>
> **Rafael**: But like some wol* not all the wolves are gonna be dead 'cause like it's not all the wolves are not in one forest I mean one place yeah one place. (Thread 2)

Chains with intervening speech. Interruption, interjection, and overlapping speech are very common in children's discussions. A chain may be interrupted at any point. If the speaker is able to continue the line of reasoning within three speaking turns after the interruption, it is considered to be a chain with intervening speech. A chain is considered to be broken if the speaker never completes the chain, despite having produced one or two links before being interrupted. In the example in the section about the branched chain, Tevon only generates one link before being interrupted by Raven and Enrique. When the teacher brings the students back to the topic a little later, he takes the floor and extends the chain that he has previously began. In some situations, students lose the floor after a strong interruption. However, if the teacher steps in and gives the floor back to the first student, the chain is likely to be developed. In the following example, Roberto opposes the killing of wolves and begins to introduce a causal relationship. However, Roberto is interrupted by Brian. The teacher stops Brian before he can express his thought and asks Roberto to continue his reasoning. Roberto then verbalizes the ecosystem-oxygen model.

Roberto: I say I say no [**B1**] because, then they gonna- I say no [**B1**] because then the, what I said that the oxygen. [**B4**]

Brian: I I say no because like there there has- [*interrupted by the teacher*]

Teacher: [talking to Roberto] Excuse us. We we need to hear your reason again. Complete answers, complete information. [*teacher's prompting*]

Roberto: I say no [**B1**] because, mm then if they do the, [**B1**] then the elk are gonna eat all the grass and trees [**B3**], and then the ox- there's there's not gonna be a, a lot of oxygen [**B4**] then the people gonna die. [**B5**]

After identifying all the causal chains in a discussion, the coder further identified the structure for each causal chain. The intercoder reliability for structure of causal chains indicated high agreement between coders (percentage of agreement was 92% and Cohen's K = .80).

3.5.3. Coding features of multilink causal links

The construction of causal chains in a group discussion is likely to be a collaborative effort rather than an individual endeavor. During the co-construction process, group members fulfill different roles in the production of co-constructed chains; these roles are to *initiate* a chain, to *add to* a chain, and to *elaborate or repeat* links in a chain. Correspondingly, there are three distinctive features of links in a multilink reasoning chain: initiated links, added links, or elaborated or repeated links. In the following example, Miguel initiates the ecosystem-oxygen model with three explicit links ("wolves eat elk", "trees produce oxygen", and "people need oxygen to survive") and two implicit links ("people kill wolves" and "elk eat trees"). Demisha adds to Miguel's chain by pointing out that "it's destroy the town." This link extends the ecosystem-oxygen model to the system level because "killing wolves" is no longer a single event but places a detrimental impact on the entire town. Eric takes the next turn to elaborate Miguel's chain, in particular, the statement that "wolves eat elk." The change in wolf population impacts the elk population. Dallas extends Eric's chain with three more links, which are actually a repetition of Miguel's chain. Taken together, four students construct this ecosystem-oxygen model, each playing a unique role. The percentage of agreement between coders for coding initiated links was 90% (Cohen's K = .83), for added links was 93% (Cohen's K = .81), and for elaborated or repeated links was 92% (Cohen's K = .79).

Miguel: No they shouldn't 'casue if the wolves don't kill the elk, then there're not gonna have trees and they're not gonna survive and 'cause they give us oxygent [oxygen]. Go on. [*Initiated links*]
Eric: They should.
Andrea: Why?
Demisha: They should.
Dallas: Why?
Demisha: Because it's destroy the town and kill other animals. [*Added links*]
Eric: But then 'cause if they do, there're gonna be more elk basically. More elk and more stuff. [*Elaborated/repeated links*]
Demisha: Shut up Eric.
Eric: I start first. Be quiet.
Teacher: One at a time please.
Dallas: If it's more elk and it's gonna it's like less trees and if it's less trees is less oxygen and if it's less oxygen it's less people. [*Elaborated/repeated links*]

3.5.4. Coding elements of multilink causal links

Segmentation. Multilink causal chains are further broken down into *reasoning units.* The most typical reasoning unit is a clause that contains a subject and a predicate; for example, "wolves are killing the livestock" or "ranching is going down." Segmentation allows speakers' repetitions and revisions to not complicate the coherence of the speaker's reasoning. For instance, "so they[wolves] are gonna start probably have some kill probably deer elk or something" is coded as one reasoning unit despite the repetitions and revisions.

Clauses that are linked by coordinating conjunctions, such as *and*, *but*, *or*, *yet*, *then*, *however*, *and then*, *so* (in the sense of consequently or therefore), and *like* (in the sense of for example or for instance), are segmented into two reasoning units. In clauses that are linked by subordinating conjunctions, segmentation is applied to a subordinate adverbial clause, but not to a subordinate noun clause or a subordinate adjective clause.

A subordinate noun clause serves as an object, subject, or complement of a linking verb in the main clause, being connected to the main clause by the subordinating conjunctions *that*, *who*, or *what*. In this case, the two clauses contribute to one meaning and so are not segmented. For example, the clause “that's [that] probably was what the wolves feel like” is coded as one reasoning unit. A subordinate adjective clause serves as adjectives of a noun using subordinating conjunctions *that, who,* and *what.* In this case, the subordinate clause is not considered a separate entity from the noun that it modifies and therefore is not segmented. For example, “wolves might kill people who are sleeping” is coded as one reasoning unit.

Subordinate adverbial clauses are often used to explain why, how, when, where, to what extent, and under what conditions. Common subordinating conjunctions include *because*, *if*, *as long/much/many as*, *since*, *that, how*, *unless*, *although*, *until*, *once*, *after*, *when*, *before*, *while*, *where*, *so* (in the sense of so that or in order to), and *like* (in the sense of such as or similar to). Since the main clause and the subordinate clause each carries a specific meaning to contribute to a united logical relation, each clause can be treated as one reasoning unit. For example, “if we kill all the wolves, the elk is gonna keep on eating the um um plants” is composed of one main clause “the elk is gonna keep on eating the um um plants” and one subordinate clause “if we kill all the wolves.” Each of the two clauses represents a simple relation, *kill* [we, wolves] and *eat* [elk, plants]. The two simple relations are connected by a causal relation in the proposition. I segmented all the causal chains and a second coder independently coded 20% of the data. The percentage of agreement between the two coders in segmentation was 98% (Cohen's K = .89).

Elements of multilink causal reasoning links. Reasoning units are categorized according to five basic features: three categories of concepts (surface feature, relational concept and system concept) and two levels of relationship (simple relation and causal relation). *Surface feature* refers to “literal objects, concepts, or entities explicitly described in a problem statement or a situation” (Chi & VanLehn, 2012, p. 178). For example, wolf, elk, and plant are surface features. *Simple relation* describes the relationships between surface features; for example, wolf and elk are connected by the simple relation *eat*. *Relational concept* refers to an inference derived from a simple relation. For example, the inference of ‘wolves eat elk’ is ‘wolf population changes’; a student described this as “there will not be that much of wolves.” *Causal relation* represents the relationship between two relational concepts. In this example, wolves eat elk and elk eat trees, so the increase of wolf population will cause a decrease in the elk population and an increase in the tree population. Therefore, the relational concepts “increase of wolf population” and “decrease of elk population” form a causal relation and the relational concepts “decrease of elk population” and “increase of tree population” form a second causal relation. *System concept* refers to a concept that applies to a larger system. For example, “a change of wolf, elk, and tree population will affect the balance of the ecosystem” is a system concept.

In multilink causal reasoning chains, surface features and simple relations are lower-order concepts; relational concepts, system concepts, and causal relations are higher-order concepts. The percentage of agreement between coders for coding surface features was 96% (Cohen's K = .96), for simple relations was 91% (Cohen's K= .89), for relational concepts was 88% (Cohen's K= .85), for causal relation was 87% (Cohen's K= .82), and for system concepts was 96% (Cohen's K= .78).

3.6 Coding Discourse Moves

Three types of discourse moves are identified in collaborative discussions and examined for how they are related to the production of multilink causal chains, including *extending moves*, *leadership moves,* and *reasoning moves*. Discourse moves are coded at the turn level (*N* = 3354).

3.6.1 Coding extending moves

Following Chiu's (2008) decision tree (see Table 3), four types of extending moves were coded: *new topic, support*, *refute*, or *neutral*. The following sections describe the coding criteria for each extending move. I coded all the extending moves in the 24 discussion transcripts and a second coder independently coded 20% of the transcripts. The percentage of agreement between the two coders was 93% (Cohen's K = .84).

Table 3
Decision Tree of Extending Moves

Does the speaker respond to the previous speaker?
- No, code as *new topic*
- Yes, does the speaker agree with the previous speaker?
    - Yes, code as *support*
    - No, does the speaker disagree with the previous speaker?
        - Yes, code as *refute*
        - No, code as *neutral*

New topic. New topic is coded when a speaker brings up a new topic without responding to the previous speaker. Four types of new topic moves are identified in the students' discussion transcripts: [1] a speaker brings up a completely new idea that has not been mentioned by the previous speaker; [2] a speaker takes the turn to invite another group member to contribute some ideas, without addressing whether he or she agrees or disagrees with the previous speaker; [3] a speaker starts a poll and requests all the group members to vote; or [4] a speaker goes off topic and brings up an irrelevant subject to the group.

Support. As suggested by Lin, Anderson et al. (2015), support moves include instances when speakers "acknowledge, justify, praise, or agreeably elaborate what the previous speaker said, or offer help based on the previous speaker's request" (p. 88). Support moves indicate agreement, acknowledgement, and intention to increase the breadth of the current argumentation. Teachers' support moves are often part of managing the discussion. For example, a teacher may prompt a student to share his/her reasoning or may ask for clarification.

Refute. This move is coded when a speaker has "made corrections, suggested alternatives, or posed challenges to the previous speaker" (Lin, Anderson et al., 2015, p. 88). Refuting moves indicate a speaker's disposition to consider the other side of an issue, the soundness of an argument, or possible alternatives to a solution. Refuting moves contribute to extending the current argumentation in depth.

Neutral. The neutral move is coded when the speaker neither supports nor refutes the previous speaker. The speaker may take the floor to manage turn taking. In this case, the speaker reminds others of the discussion rules, asks clarifying questions or makes procedural requests. In a second case, the speaker responds to procedural requests, such as "can you come back to me?" or "tell me which expert group you are in." In a third case, a student or the teacher starts a poll and the whole group take turns to state their position. The turn that starts the poll is coded as a *new topic*, and the turns that respond to the poll are coded as *neutral.* Here is an example of coding extending moves in a discussion segment.

Jonathan	They shouldn't die because then the wolves' population is going down.	[New topic]
Lucia	Well they will be back if the population goes down.	[Refute]
Jonathan	Because then the prices will go high in a way.	[Support]
Alondra	The elk the elk eat the trees.	[New topic]
Lucia	No elk stay the same. Actually if the wolves don't get killed, the price will go up.	[Refute]
Jessica	So do you still think they should kill them?	[Neutral]
Lucia	Mhm.	[Neutral]
Jonathan	But the trees go down.	[Refute]
Natalia	Yeah 'cause if we kill the wolves, the elks are the elks going (up) the aspen trees going down.	[Support]
Jessica	If the trees go down, then there will be less less oxygen. Less and less.	[Support]

3.6.2 Coding leadership moves

The present study also identifies five types of leadership moves based on Li et al.'s (2007) categorization, including *turn management*, *argument development*, *planning and organizing*, *clarifying and revoicing*, and *topic control.* Notably, this study replaced the category "acknowledgement" in Li et al.'s with a new category called 'clarifying and revoicing' because acknowledgement moves were very rare in the present dataset (five speaking turns including one student turn and four teacher turns), whereas students frequently asked each other to clarify their thoughts through repeating and rephrasing.

Specifically*, turn management* moves refer to the actions that students take to solicit opinions from group members, to arrange the order of turn-taking, to intervene when interruption occurs, or to help others gain or hold the floor. *Argument development* moves refers to the actions by which a speaker solicits opinions, reasons and evidence from group members, usually in the format of why and how questions. *Planning and organizing* moves include monitoring discussion process, summarizing opinions, or starting a poll to collect group members' viewpoints on the central question. *Clarifying and revoicing* moves refer to the action of restating what other just said or asking for clarification questions so that the group maintains a shared understanding among group members. *Topic control* moves refer to the speaker's effort to redirecting the topic back to the central question or bringing up the other side of the issue. The coding criteria and examples, along with coding reliabilities are presented in Table 4.

Table 4

The Coding Scheme of Leadership Moves

Coding Category	Definition	Example	Cohen's Kappa
Turn management	Arranging the order of turn-taking; intervening when interruption occurs; helping others gain or hold the floor.	- Jonathan, what do you think? - Let Jonathan say something! - We will start with Jonathan.	0.89

Table 4 (cont.)

Coding Category	Definition	Example	Cohen's Kappa
Argument development	Soliciting opinions, reasons and evidence from group members	- Why shouldn't they kill the wolves? - That you want the wolves to be killed? How was that? Explain to us. - But how did they know it was a wolf?	0.78
Planning and organizing	Monitoring discussion process, summarizing opinions, starting a poll to collect group members' viewpoints on the central question.	- Thumbs up if you guys think they should kill them, and thumbs down if you think they shouldn't kill them. - We need to go man. We need to go. She said three minutes. We passed three minutes.	0.73
Clarifying and revoicing	Asking for clarification, explanation, or justification; restating the question or the argument.	- You mean they kill the the they kill the children? - Will you say it again? Repeat it. - What do you mean with that you didn't want that much?	0.80
Topic control	Redirecting the topic back to the central question or bringing up the other side of the issue.	- So let's go back to the big question, guys. - Well, now let's put ourselves in the townspeople's position.	0.80

3.6.3 Coding reasoning moves

Four types of reasoning moves were coded in the transcripts: *positioning and justifying for wolf hunting, evaluating the credibility of evidence, proposing alternate solutions*, and *other topics.* The coding criteria and examples and coding reliabilities are presented in Table 5.

As a direct response to the central question, students take a position on whether or not the townspeople should be allowed to hire professional hunters to kill the wolves and then *justify* their position based on supporting evidence from the text or other reliable sources, or through predicting what is likely to happen if wolves are killed or not. Students also place themselves or others in the scenario to envision what they would do if they were wolves or the townspeople. By immersing oneself or others in the story world, students are able to justify their position based on moral principles such as fairness, empathy, or common good.

However, indirect responses to the central question were also observed in the discussion. One type of indirect response is to *evaluate the credibility of evidence*, as commonly seen in legal reasoning. Students questioned if the townspeople had really seen a pack of wolves in the town or nearby or if any of the reported wolf attack was confirmed. If wolves did not come to the town or did not cause any harm to the people, making a decision on eradicating wolves would be unjust and unfair. Another type of indirect response is to *propose an alternate solution* to the wolf problem so that students do not have to make a decision on whether or not to kill wolves. Four most popular alternate solutions were: building a fence to separate wolves and people, moving wolves to another place, moving townspeople to another place, and domesticating the wolves. Alternate solutions are usually based on children's naïve understanding of wolves and often contain incorrect information such as wolves can be tamed like dogs.

This study also identifies *other topics* that stray away from the central question, which include imported wolf knowledge and off-topic talk. In collaborative discussions, children often share information through knowledge telling based on what they have read or heard from outside sources. Knowledge telling is a way to generate topic-relevant ideas based on free association (Bereiter & Scardamalia, 1987). This strategy is commonly used in writing, especially among immature writers. When given a topic to discussion or write, learners first identify key words that are related to the main topic and then use them as cues to search for associated information in memory. The search favors the most accessible information, and often lacks synthesis of knowledge or account of coherence (Bereiter & Scardamalia, 1987). The wolf-related information generated from knowledge telling usually has no direct relation with the central question. Therefore, imported wolf knowledge and off-topic talk are categorized as *other topics*.

Table 5

The Coding Scheme of Reasoning Moves

Coding category	Definition	Example	Cohen's kappa
Positioning and justifying for wolf hunting			0.90
Positioning only	Take a position on the central question without providing any reasons or explanation.	- I think wolves shouldn't be killed. - I agree/disagree with Juan that wolves should be killed.	
Positioning based on text information	Justify one's position with facts, text materials, or relevant information from other resources such as books, films, or TV shows.	- I think wolves shouldn't be killed because the wolves ain't the only animal that's killing all the sheep and stuff. - I agree with the three of you 'cause in the booklet that I read, it said that nowhere in the United States there has been a wolf that has killed a human human.	
Positioning based on hypotheses	Predict what is likely to happen if wolves are killed or not, or place oneself or others in the situation to infer the character's reactions.	- I think they should uh stay alive, because the like if the wolves eat the elk, there will be more tree. - If I was , if ok, if I was in the townspeople's position, I I would still, I think I will still not want the wolves to be killed because they're not harming anybody.	
Positioning based on moral principles	Justify one's position based on one's own or the character's emotional reaction to the problem, or based on moral principles.	- Because they[wolves] have as much right as us to be there because that's more of their natural habitat. - I think they shouldn't kill the wolves because um like the wolves have feelings and if they kill you, you'll feel sad.	
Evaluating the credibility of evidence			0.82
Questioning the seeing of wolves	- And that's what they anyways they will not come because they are shy of people. In the video it said they are shy, so why would they come if they are shy of us?		
Questioning reported wolf attacks	- They'll just heard the wolves howl howlin' around the dog. They claiming a wolves killed it, but they don't know if the wolves kill for sure.		

Table 5 (cont.)

Coding category	Example	Cohen's kappa
Proposing alternate solutions		0.95
Building a fence to separate wolves and people	- Why don't they put a fence right here where where elks can live right here. And right here is for the wolves. And right here is for the the hunters. - I'm talkin' about a wooden fence but it's got a like a sumpin' really strong right behind it.	
Moving wolves to a different place	- I think maybe that they should use special equipment to tase them make them go to sleep and then like move them to a different area um away from the people.	
Moving townspeople to a different place	- The townspeople. The townspeople? They need to move. - How are they gonna move? They they have to break all the houses and they have to pay for houses and probably they don't have some money.	
Domesticating/ taming/capturing the wolves	- I think that uh like a wolf biologist could like take, like a two week-old cub and then like take it to like a day care. Take it till it gets really big and then let it go, and let it be wild, so then it won't be as dangerous, to the people.	
Other topics		0.85
Wolf knowledge telling	- I read that book and it says wolves can run very fast. - A coyote is a wolf's cousin. - There's an alpha female and an alpha male. There's only one of each in the pack.	
Off-topic talk	- I got stuff on my shirt. - Like sometimes when when well when I went in Mexico, my I think my uncle or something we we start running we start running some sheep.	

CHAPTER 4

RESULTS

4.1 Pretest Performance

Separate two-level regression analyses were conducted to check whether there was a difference between the three expert groups (i.e., ecosystem, economy, or public policy) in pretest reading comprehension or oral English fluency (measured by the rapid automatized naming task). Fixed effects include gender, ethnicity, expert groups, and whether or not the student was enrolled in the Individualized Education Program (IEP). The data have a nested structure—students nested within groups ($N = 24$) and groups nested within classrooms ($N = 12$). Three-level regression models were constructed first but abandoned later because the classroom-level difference in pretest performance was not significant. Therefore, the final model collapsed three levels into two—the groups were entered at the second level to account for cohort effects.

The analyses of reading comprehension indicated that girls tended to have higher reading scores than boys but the difference was marginal, $F(1, 130) = 3.68$, $p = .057$. IEP students had lower reading scores than non-IEP students, $F(1, 130) = 4.59$, $p = .034$. Three of the expert groups were comparable in reading comprehension, $F(2, 130) = 1.09$, $p = .34$. No ethnicity difference was found in reading comprehension, $F(2, 130) = 1.83$, $p = .16$. Between-group difference was significant, $Z = 1.83$, $p = .033$. About 18% of the variance was due to cohort effects.

The same two-level model was created to examine differences in object naming task, which is an indicator of oral English fluency. The results indicated that African American students had better oral English fluency than Latino/a children, $F(1, 130) = 20.79$, $p < .01$. No significant difference was found between boys and girls, among the three expert groups, or between IEP and non-IEP students. Group differences were not significant, $Z = 1.07$, $p = .14$.

Individual social characteristics. Seven measures of individual social characteristics are derived from peer nominations in the pre-intervention sociometric questionnaire, including talkativeness, good idea nomination, leadership nomination, peer liking rating, degree centrality, betweenness centrality, and information centrality. The correlations among these variables are presented in Table 6. Seven separate two-level regression analyses were performed to evaluate individual differences in social characteristics with gender, ethnicity, expert groups, and whether or not a child enrolled in IEP as fixed effects and group as a random variable. Girls received more good idea nominations, $F(1, 130) = 7.20$, $p = .008$, more leader nominations, $F(1, 130) = 5.06$, $p = .026$, and higher peer liking rating than boys, $F(1, 130) = 13.52$, $p < .001$. IEP students were less likely to be nominated as 'having good ideas' as compared to non-IEP students, $F(1, 130) = 5.17$, $p = .025$. Ecosystem students were more talkative, $F(1, 130) = 6.60$, $p = .011$, more likely to be nominated as 'having good ideas', $F(1, 130) = 5.06$, $p = .026$, and more liked by their classmates, $F(1, 130) = 6.00$, $p = .016$, as compared to economy students. No difference was found in any of the centrality indices between the ecosystem group and the public policy group except that ecosystem students showed higher information centrality than the public policy students, $F(1, 130) = 6.20$, $p = .014$. No between-group difference was observed.

Table 6

Correlations of Seven Social Characteristics

	1	2	3	4	5	6	7
1. Talkativeness	-						
2. Good idea nomination	$.32^{***}$	-					
3. Leader nomination	$.16^{*}$	$.75^{***}$	-				
4. Peer liking rating	$.17^{*}$	$.47^{***}$	$.44^{***}$	-			
5. Degree centrality	$.25^{**}$	$.39^{***}$	$.30^{***}$	$.49^{***}$	-		
6. Betweenness centrality	$.26^{**}$	$.16^{*}$	.15	$.25^{**}$	$.48^{***}$	-	
7. Information centrality	$.24^{**}$	$.26^{**}$	.12	$.45^{***}$	$.73^{***}$	$.52^{***}$	-

Note. Significance level: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

4.2 Analyses of Features and Structures of Multilink Causal Reasoning

4.2.1. Descriptive statistics of chain production

A total number of 122 multilink causal chains were identified in 24 group discussions, resulting in average 5.08 multilink causal chains per discussion with a range of 0–11 causal chains (discussion time ranged between 7 and 25 minutes). These chains were generated by 92 (58%) students, and the remained 68 (42%) students did not contribute to any chains. Among the 92 students, 42 (46%) students contributed to one causal chain, 27 (29%) students contributed to two causal chains, seven (8%) students contributed to three causal chains, 13 (14%) students contributed to four chains, and the remaining three students participated in the production of five, six, and nine chains, respectively. Seventy-four (61%) chains were generated by a single child and 48 (39%) chains were constructed by a group of 2-6 students. Most teachers did not participate in the production of causal chains, except for one teacher who prompted the production of two links in an eight-link chain and contributed one link herself. At the group level, 12 groups produced five or more chains, ten groups produced one to four chains, and two groups did not produce any chains.

About 7.1% of the speaking turns (238 of 3,354) were contributed to the production of multilink causal chains. The four most frequent types were the ecosystem-oxygen model (n = 39), supply model (n = 26), tourism model (n = 16), and competition model (n = 9). The other three types were rather infrequent—there were three compensation chains, five leftover chains, and seven ecosystem-economy chains. Seventeen novel chains were identified. Eighty-three (68%) chains were straight chains, 22 (18%) chains were branched chains, 16 (13%) chains were generated after teachers' or group members' interruption, and one chain was threaded by another line of conversation.

The length of the causal chains ranged from three links to eight links. There were 49 (40%) three-link chains, 39 (32%) four-link chains, 26 (21%) five-link chains, five (4%) six-link chains, one seven-link chain, and two eight-link chains. Chains with five or more links were 2.67 times more likely to be generated by a group rather than by an individual. A chi-square association test showed that there was a significant association between length of causal chain (3, 4, 5 or more links) and number of contributors (1, 2, 3 or more students), $\chi^2(4) = 12.60$, $p = .013$. As anticipated, longer chains usually take more than one speaking turn to generate. Chains with five or more links was 3.74 times more likely to take four or more speaking turns to produce than taking only one speaking turn and 2.42 times more likely to take two or three speaking turns than taking one speaking turn. The chi-square test indicated a significant association between length of causal chain and number of speaking turns to complete a chain, $\chi^2(4) = 15.90$, $p = .003$.

4.2.2. The influence of group features on chain production

Among the 24 collaborative groups, 12 groups contained large number of Latino/a children and the other 12 groups contained large number of African American children. Nine group features were examined in this study: group average reading comprehension; group average oral English fluency; group average talkativeness, good idea rating, leadership rating, and peer liking rating; and the average score of three indicators of group members' social status in the class, namely, degree centrality, betweenness centrality, and information centrality. Group-level social characteristics are calculated based on the Z score of group members' individual scores, which represents each group member's relative standing in the class. For instance, high group-level talkativeness indicates that this group gathered more talkative children in the class.

On average, groups with dominant African American children had better oral English fluency (AA: *mean* = 64.9, *SD* = 7.3; Latino/a: *mean* = 53.7, *SD* = 6.6) and higher reading ability

(African American: *mean* = 18.0, *SD* = 7.1; Latino/a: *mean* = 13.7, *SD* = 3.7), as compared to groups with dominant Latino/a children. A correlation analysis was conducted to calculate the partial Spearman correlations of the nine group-level social and cognitive features and the five measures of multilink causal reasoning in each group. The five outcome measures include number of different type of causal chains, total number of causal chains, total number of causal links, total number of single-person causal chains, and total number of co-constructed chains. The correlation analysis controlled for group size, duration of talk, and total number of speaking turns in the discussion. The results were presented in Table 7.

Table 7

Spearman Partial Correlations of Group-Level Social and Cognitive Attributes and Measures of Multilink Causal Reasoning (N=24)

	Types of Causal chains	Number of causal chains	Number of causal links	Number of single-person chains	Number of co-constructed chains
1 Reading comprehension	.55**	.43*	.44*	.41	.35
2 Basic English proficiency	.03	.08	.08	.09	.07
3 Talkativeness	.23	.52*	.47*	.59**	.32
4 Good idea rating	.20	.07	.05	.11	.07
5 Leadership rating	.59**	.37	.33	.33	.42
6 Peer liking rating	-.06	-.29	-.34	-.26	-.27
7 Degree centrality	-.24	-.21	-.17	-.12	-.30
8 Betweenness centrality	.06	.19	.17	.28	.10
9 Information centrality	-.49*	-.41	-.40	-.32	-.44*

Note. The scores of nine social and cognitive features are group means. * $p < .05$, ** $p < .01$.

According to Table 7, group average reading comprehension showed a significantly positive correlation with the types of causal chains (r_s = .55) and the number of causal chains (r_s = .44), suggesting that groups that contained more good readers tended to produce more multilink reasoning chains that groups that contained more poor readers. Talkativeness was also a significant predictor of multilink reasoning (r_s = .52). Groups with more talkative children tended to produce more single-person causal chains (r_s = .59) than co-constructed chains. Group average leadership rating had a highly positive correlation (r_s = .59) with types of causal chains and a moderately positive correlation (r_s = .42) with number of co-constructed chains. This result suggests that groups with more leaders tend to generate more types of chains and with these leaders in the group, students were more likely to co-construct chains together.

However, group average oral English fluency showed a low correlation (r_s = .08) with the number of multilink causal chains produced by the group, suggesting that oral English fluency was not a critical factor for producing multilink causal chains at the group level. Students with high oral English fluency may help those who had difficulty expressing themselves to participate in chain construction. Another result was that group average information centrality, an indicator of group members' social status in the class, had a significantly negative correlation with types of causal chains (r_s = -.49) and the number of co-constructed causal chains (r_s = -.44). Previous analysis of children's social characteristics showed that ecosystem students had higher information centrality than students from the other two expert groups. The negative relationship between group-level information centrality and number of co-constructed chains produced by a group suggests that ecosystem students were less likely to co-construct causal chain with group members as compared to economy and public policy students.

4.2.3 The effect of individual characteristics on production of causal chains

Production of causal chains. Children's multilink causal reasoning was evaluated by three measures: types of causal chains, total number of causal chains, and total number of causal links. Table 8 summarizes the means and standard deviations of the three measures by gender, ethnicity and expert groups. Three two-level Poisson regression analyses were conducted using gender, ethnicity, expert group (ecosystem, economy or public policy), and whether or not students enrolled in the Individualized Education Program (IEP) as fixed effects and group as a random effect. The covariates include reading comprehension, oral English fluency, talkativeness, good idea nominations, leader nominations, information centrality, and ratings of peer liking. Degree centrality and betweenness centrality were first included but dropped later because they did not predict any of the outcome measures.

Individual talkativeness, oral English fluency, and peer liking rating significantly predicted the *number of causal chains*, $p < .05$. Economy students produced more causal chains than the ecosystem students, $F(1, 123) = 8.18$, $p = .005$, and the public policy students, $F(1, 123) = 10.69$, $p < .001$. There was no difference between the ecosystem students and the public policy students, $F(1, 123) = 0.59$, $p = .44$. No other predictors were significant.

There were significant differences in the *number of causal links* among three ethnic groups (Latino/a, African American, and other) and among three expert groups. Latino/a children produced more links than the other two ethnic groups and economy students produced more links than ecosystem and public policy students. An interaction of ethnicity and expert group was added to the model and found significant—Latino/a children showed similar performance across three expert groups; however, African American children in ecosystem and public policy groups showed significantly lower performance than students in the economy group. Economy students also generated more *types of causal chains* than ecosystem students, $F(1, 123) = 4.12$, $p = .045$, and public policy students, $F(1, 123) = 5.29$, $p = .023$. No difference was found between ecosystem and public policy students in terms of types of causal chains. Oral English fluency and talkativeness both significantly predicted the number of causal links and the variety of causal chains, *ps* < .05. Students who had higher reading ability, $F(1, 119) = 4.37$, $p = .039$, and students who received more leader nominations produced more causal links, $F(1, 119) = 4.23$, $p = .042$. No other covariates were found significant.

Table 8

Means and Standard Deviations of Measures of Multilink Causal Reasoning by Gender, Ethnicity, and Expert Groups

	Gender		**Ethnicity**			**Expert groups**		
	Girls ($N = 93$)	Boys ($N = 67$)	Latino/a ($N = 73$)	African American ($N = 70$)	Other ($N = 17$)	Ecosystem ($N = 60$)	Economy ($N = 50$)	Public policy ($N = 50$)
Production of causal chains								
Types of causal chains	1.01 (1.27)	0.93 (0.99)	1.00 (0.93)	0.97 (1.43)	0.88 (0.78)	1.08 (1.18)	1.10 (1.38)	0.72 (0.83)
Number of causal chains	1.16 (1.53)	1.19 (1.38)	1.25 (1.30)	1.14 (1.71)	1.00 (1.00)	1.32 (1.50)	1.40 (1.76)	0.78 (0.95)
Number of causal links	3.48 (5.17)	3.49 (4.29)	3.82 (4.27)	3.39 (5.62)	2.47 (3.10)	4.00 (4.96)	4.10 (5.72)	2.26 (3.23)
Features of causal links								
Initiated links	2.09 (3.57)	2.04 (2.72)	2.23 (2.96)	1.97 (3.62)	1.76 (2.75)	2.33 (3.07)	2.68 (4.08)	1.14 (2.15)
Added links	0.57 (1.20)	0.64 (1.15)	0.63 (1.12)	0.63 (1.34)	0.35 (0.49)	0.70 (1.37)	0.66 (1.26)	0.42 (0.78)
Elaborated/repeated links	0.83 (1.52)	0.81 (1.40)	0.96 (1.48)	0.79 (1.59)	0.35 (0.61)	0.97 (1.72)	0.76 (1.44)	0.70 (1.15)
Elements of causal links								
Lower-order concepts	5.38 (7.53)	5.63 (7.03)	6.01 (6.81)	5.11 (8.24)	4.71 (5.11)	6.47 (7.60)	6.62 (8.59)	3.16 (4.64)
Surface features	2.89 (4.33)	2.93 (3.67)	3.18 (3.68)	2.64 (4.62)	2.82 (3.11)	3.47 (4.26)	3.58 (4.82)	1.56 (2.36)
Simple relations	2.48 (3.40)	2.70 (3.62)	2.84 (3.31)	2.47 (3.90)	1.88 (2.15)	3.00 (3.62)	3.04 (4.00)	1.60 (2.45)
Higher-order concepts	4.61 (7.71)	4.73 (8.90)	4.65 (5.55)	4.84 (8.68)	3.94 (4.62)	5.92 (7.69)	5.18 (8.06)	2.63 (4.02)
Relational concepts	2.31 (4.29)	2.49 (3.34)	2.29 (2.97)	2.61 (4.96)	1.88 (2.29)	3.03 (4.48)	2.56 (4.27)	1.44 (2.39)
Causal relations	2.08 (3.33)	2.15 (2.65)	2.20 (2.65)	2.13 (3.62)	1.65 (2.13)	2.60 (3.18)	2.50 (3.74)	1.13 (1.64)
System concepts	0.22 (0.64)	0.09 (0.45)	0.16 (0.55)	0.10 (0.42)	0.41 (1.00)	0.28 (0.76)	0.12 (0.52)	0.06 (0.24)

Features of multilink causal links. Three types of links were identified during the construction of multilink reasoning chains—initiated links, added links, and elaborated or repeated links. Table 8 presented the means and standard deviations of three features of causal links by gender, ethnicity, and expert groups. Three two-level Poisson regression analyses were conducted to investigate the effects of individual characteristics on the production of initiated links, added links, and elaborated/repeated links. The same set of fixed effects and covariates were entered in the model.

Initiated links. There was a significant ethnicity difference in the number of initiated links produced by each child, $F(2, 123) = 6.01$, $p = .003$, and significant differences among three expert groups, $F(2, 123) = 21.24$, $p < .001$. Latino/a children were more likely to initiate a causal chain than African American children, $F(1, 123) = 11.16$, $p = .001$. Economy students were more likely to initiate a chain of reasoning, as compared to ecosystem students, $F(1, 123) = 30.58$, $p < .001$, and public policy students, $F(1, 123) = 27.56$, $p < .001$, while there was no difference between the ecosystem group and the public policy group. An ethnicity by expert group interaction was added to the model but was found not significant, $F(1, 119) = 2.12$, $p = .082$. Students who had better oral English fluency and who were more talkative were more likely to initiate a chain, $\beta_{English} = 0.03$, $F(1, 119) = 28.62$, $p < .001$, and $\beta_{talkative} = 0.75$, $F(1, 119) = 75.65$, $p < .001$. No other predictors were significant.

Added links. There was no ethnicity difference in terms of number of added links produced by each individual, but there were significant differences among three expert groups. Economy students were more likely to add to a chain than ecosystem students, $F(1, 123) = 5.11$, $p = .026$, but only had a marginal difference from the public policy students, $F(1, 123) = 3.17$, $p = .078$. Students who had better oral English fluency ($\beta_{English} = 0.03$) and who were more talkative ($\beta_{talkative}$ = 0.61) were more likely to add to a chain, *ps* < .001. Children's information centrality showed a significantly negative effect on the likelihood of adding to a chain, $\beta_{centrality}$ = -0.34, $F(1, 123) = 4.17$, p =.043. This result indicated that students who were more socially centered in the class were less likely to add to a chain. Previous analysis of individual social characteristics indicated that ecosystem students showed higher information centrality than economy and public policy students. Thus, this finding implies that ecosystem students may be less likely to add to a chain as compared to economy and public policy students.

Elaborated or repeated links. Latino/a children were more likely to elaborate or repeat links than African American children, $F(1, 123) = 3.88$, $p = .051$. The difference among three expert groups in elaborating or repeating links was not significant, $F(2, 123) = 0.86$, $p = .43$. A notable result is that oral English fluency did not predict the likelihood that a child elaborated or repeated causal links, $F(1, 123) = 2.08$, $p = .15$, which suggests that children who have low oral English fluency probably participated in chain construction through elaborating or repeating links in chains created by their peers who were more competent to generate new ideas. Students who were more talkative and who received higher peer liking rating were more likely to elaborate or repeat causal links, $\beta_{talkative}$ = 0.42, $F(1, 123) = 11.66$, $p = .001$, and $\beta_{peer\ liking}$ = 0.39, $F(1, 123) = 6.88$, p =.010; whereas students with high information centrality were less likely to elaborate or repeat causal links, $\beta_{centrality}$ = -0.30, $F(1, 123) = 4.62$, p =.034. No other predictors were significant.

Elements of multilink causal links. Five elements of multilink causal reasoning were identified in this study—surface features, simple relations, relational concepts, causal relations, and system concepts. The first two features were grouped as lower-order concepts and the last three features were grouped as higher-order concepts. Two-level Poisson regression analysis was conducted to examine the effect of individual characteristics on the production of lower-order and higher-order concepts. The same set of fixed effects, random effect and covariates were entered into the model. The means and standard deviations of lower- and higher-order concepts by gender, ethnicity, and expert groups were presented in Table 8.

Lower-order concepts. There was a significant ethnicity difference, $F(2, 119) = 9.84$, $p < .001$, a significant difference among three expert groups, $F(2, 119) = 20.70$, $p < .001$, and a significant ethnicity by expert group interaction, $F(4, 119) = 111.74$, $p < .001$. African American children in the ecosystem and the public policy groups generated significantly fewer lower-order concepts than the other groups, but African American children in the economy groups generated most lower-order concepts as compared to all the other groups. Boys tended to generate more lower-order concepts than girls, $F(1, 119) = 2.87$, $p = .093$. Four individual characteristics significantly predicted children's production of lower-order concepts, including oral English fluency, $\beta = 0.03$, $F(1, 119) = 53.33$, $p < .001$, talkativeness, $\beta = 0.75$, $F(1, 119) = 203.47$, $p < .001$, leader nominations, $\beta = 0.17$, $F(1, 119) = 7.45$, $p = .007$, and information centrality, $\beta = 0.13$, $F(1, 119) = 4.50$, $p = .036$.

Higher-order concepts. Boys generated more higher-order concepts than girls, $F(1, 119) = 7.52$, $p = .007$. There was also a significant ethnicity difference, $F(2, 119) = 9.40$, $p < .001$, a significant difference among three expert groups, $F(2, 119) = 10.76$, $p < .001$, and a significant ethnicity by expert group interaction, $F(4, 119) = 5.26$, $p < .001$. African American children in the economy groups produced most higher-order concepts as compared to all the other groups. Three individual characteristics significantly predicted children's production of higher-order concepts, including oral English fluency, $\beta = 0.03$, $F(1, 119) = 53.65$, $p < .001$, talkativeness, $\beta = 0.69$, $F(1, 119) = 138.71$, $p < .001$, and leader nominations, $\beta = 0.19$, $F(1, 119) = 8.01$, $p = .006$.

To summarize, three major findings can be derived from the individual-level analyses: [1] economy students outperformed ecosystem and public policy students in the production of multilink causal chains; [2] children's oral English fluency and talkativeness significantly predicted their multilink causal reasoning; and [3] socially and cognitively advanced children were more likely to initiate and add to a chain, while socially and cognitively less advanced children were more likely to elaborate or repeat links in a chain.

4.2.4 Analyzing the length of multilink causal chains

Since 40% of the causal chains were generated by a group of students rather than a single child, the analysis of length of multilink causal chains was conducted at the chain level. There were 122 causal chains in total, generated by 92 students. The social and cognitive features of contributors were used as covariates in the analysis. For chains generated by an individual, the child's reading comprehension, oral English fluency, and social characteristics were used. For chains generated by a group, the average score of social and cognitive measures of all the contributors was used.

The length of the causal chain ranges from three to eight links, which were grouped into four categories: three links (N = 49), four links (N =39), five links (N = 26), and six to eight links (N = 8). Length of causal chain was analyzed using multinomial logistic regression with the following variables as predictors: number of contributors (1, 2, 3 or more), number of speaking turns that are used to generate the chain (1, 2, 3 or more), and contributors' oral English fluency, reading comprehension, talkativeness, idea nominations, leader nominations, information centrality, and peer liking rating. Following the parsimony principle, the final model kept all the significant predictors and dropped the non-significant predictors.

The final model contained four significant predictors: number of contributors (1 is used as the reference group), contributors' average oral English fluency, contributors' average information centrality, and contributors' average peer liking rating. As compared to single-person chains, five-link chains were more likely to be generated by two children, *odds ratio* $_{(2\text{ vs.}1)}$ = 3.80, χ^2 (1) = 5.18, p = .023, and three or more children, *odds ratio* $_{(\geq 3\text{ vs.}1)}$ = 28.28, χ^2 (1) = 5.97, p = .015, than by one child. Six- to eight-link chains were also more likely to be generated by two children, *odds ratio* $_{(2\text{ vs.}1)}$ = 4.69, χ^2 (1) = 3.82, p = .050, and three or more children, *odds ratio* $_{(\geq 3\text{ vs.}1)}$ = 330.96, χ^2 (1) = 10.42, p = .001, than by one child. Contributor's information centrality had a significant effect on the production of six- to eight-link chains, *odds ratio* = 8.05, χ^2 (1) = 3.94, p = .047, while oral English fluency, χ^2 (1) = 3.76, p = .053, and peer liking rating had a marginal effect on the production of six- to eight-link chains, χ^2 (1) = 3.05, p = .081. Overall, this analysis indicated that long causal chains (chains that contain five or more links) took more students and more speaking turns to generate and students who participated in the construction of long causal chains tended to have higher social and cognitive competency.

4.3 Temporal Development of Multilink Causal Reasoning

A temporal analysis was conducted to investigate the time interval between pairs of successive chains. The time interval for the first chain was calculated as the duration from the beginning of the discussion to the time when the first chain was started. The time interval for the successive chains was calculated by subtracting the start time of the previous chain from the start time of the current chain. Generalized gamma mixed regression analysis was used with time interval (seconds) as the outcome variable, and the ordinal position of each causal chain as the explanatory variable. Gamma distribution is assumed because the occurrence of causal chains during each time period has a mean of one and unknown variance. Other predictors include length of causal chain, number of contributors and number of speaking turns for producing each causal chain. Group was entered as a random variable. Chains beyond 8th was not included in the analysis due to low frequency. The results showed a significant negative relationship between ordinal position and time interval, β= -0.11, $t(85)$ = -2.17, p = .033. This result suggested that the time interval between successive pairs of causal chain decreased as the ordinal position increased. Figure 9 presented the average time interval between pairs of successive causal chains.

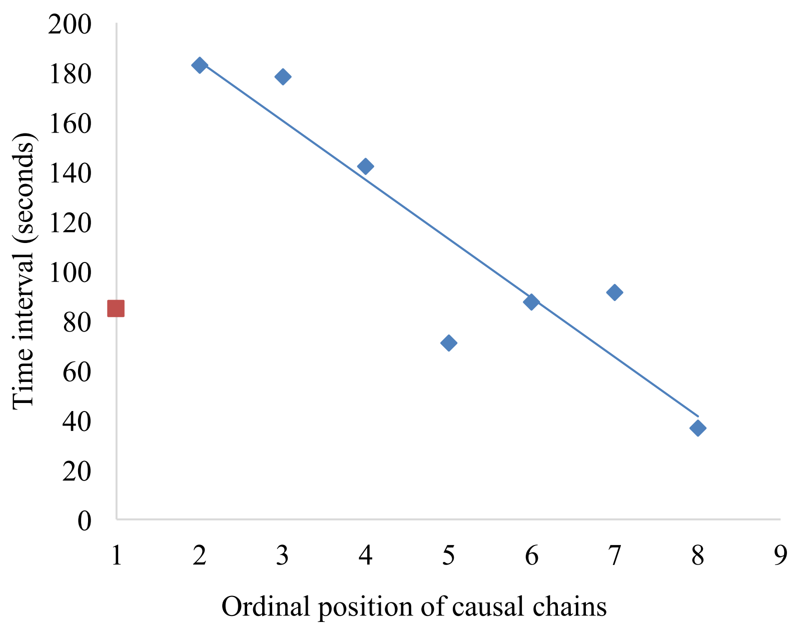

Figure 9. Average time interval between successive pairs of causal chains.

According to Figure 9, the time interval between two adjacent causal chains showed a linear decline from the second chain to the eighth chain. Number of contributors had a significantly positive effect on the time interval, β= 0.51, $t(85)$ = 2.11, p = .038, which suggested that the time interval of co-constructed chains was longer than the time interval of single-person chains. Number of speaking turns showed a marginal effect on time interval, β = -0.26, $t(85)$ = -1.65, p = .010, suggesting that the time interval between successive chains tended to decrease as the number of speaking turns increased. In other words, the more speaking turns that a group used to generate a causal chain, the more quickly a group produced a subsequent chain. Length of chain did not show a significant relationship with the time interval between pairs of chains, $\beta = 0.10$, $t(85) = 1.03$, $p = .030$, which suggested that the time interval between two short chains was comparable with the time interval between two long chains.

In addition to generalized gamma regression analysis, an event history analysis was conducted to investigate how number of contributors, number of speaking turns for producing causal chains, and length of causal chain affected the rate of occurrence of causal chains. All the discussions were videotaped from the beginning, so the problem of left censoring does not exist. Twenty-two discussions encountered different types of right censoring; for example, some teachers had to stop the discussion due to time constraints, some teachers initiated a poll by the end of the discussion to ask students to restate their position, and some teachers asked the group to take several minutes to wrap up. However, students may be able to produce more chains if they were given more time to discuss the wolf problem.

Prentice-Williams-Peterson conditional risk set model (Prentice, Williams, & Peterson, 1981) was used to analyze the time interval between recurring events during a censored period of time. In order to fit the conditional risk set model, time intervals were recalculated as the duration between the end time of the previous event and the end time of the current event. The time interval for the first event was calculated as the duration from the beginning of the discussion to the end time of this event. The full time course of the recurrent event process was modeled. For example, if a group produced one causal chain in the fifth minute during a ten-minute discussion, the first time interval was 0-5 minute with 1 event occurrence and no censoring, while the second time interval was 5-10 minute with 0 event occurrence and right censoring. The occurrence of causal chains was clustered by group and stratified by the order of each time interval. All 24 discussions were analyzed, including the two discussions that contained no causal chains. A total number of 144 time intervals were identified, during which 122 events (causal chains) were observed. The total interval time was 23,346 seconds (*mean* = 162, *SD* = 195). Three covariates were length of causal chain, number of contributors, and number of speaking turns that were used to generate the causal chain.

The global null hypothesis test based on the robust sandwich estimate indicated that the conditional model was significant, *Wald* χ^2 (3) = 38.75, $p < .001$. Length of causal chain showed a positive effect, *hazard ratio* = 1.417, χ^2 (1) = 27.40, $p < .001$. The hazard ratio estimate of length of causal chain indicated that as the discussion progressed, the rate of events (causal chains) increased by 41.7% for every one link increase in the length, meaning that long causal chains were more likely to be generated in the late stage of discussion. Number of speaking turns showed a negative relationship with time interval, *hazard ratio* = 0.849, χ^2 (1) = 4.95, $p = .026$. The hazard ration estimate of number of speaking turns suggested that as the discussion progressed, the rate of occurrence of causal chains increased by 15.1% for every one decrease in number of speaking turns. This result suggested that it took fewer and fewer speaking turns to generate a causal chain over the course of a discussion. Number of contributors did not show a significant effect, *hazard ratio* = 1.169, χ^2 (1) = 0.84, $p = .36$.

4.4 Statistical Discourse Analyses of Turn-by-Turn Chain Production

4.4.1 The proximal effects of extending moves

A statistical discourse analysis (SDA, Chiu, 2008) was performed to examine the effect of peer interaction on production of multilink causal chains in collaborative discussions. The SDA analysis targeted at the 3161 student turns that contained at least one complete idea. Teacher turns (N = 193) were not included in this analysis. Twenty-nine student turns were also removed from the analysis because the speaker's identity could not be verified and thus the concrete information of peer interaction was unable to collect at this turn. The final SDA model included 3132 identifiable student turns from 24 collaborative discussions.

To investigate the antecedent social and cognitive effects on children's production of multilink causal chains, the 3,132 turns of speaking were fitted with a two-level SDA model in which speaking turns were nested within groups. Intercepts were allowed to vary at the group level. The dependent variable was the number of causal links produced by a child at the current speaking turn (Lag-0), which ranged from zero to eight. Poisson regression analysis was used because the outcome variable followed a Poisson distribution. A primary SDA model was constructed with number of causal links produced at Lag-1, Lag-2, Lag-3, and Lag-4 as explanatory variables to explore the optimal number of lag variables to be considered in the final analysis. The results indicated that only antecedent effects at Lag-1 and Lag-2 significantly predicted the production of causal links at Lag-0, *ps* < .001. Therefore, the final model only included Lag-0, Lag-1, and Lag-2 variables. The results of SDA analysis are presented in Table 9. Non-significant variables were removed from the model following the rule of parsimony.

In Model 1, current speaker's individual characteristics were entered as explanatory variables to test if certain social and cognitive characteristics affected the development of multilink causal reasoning, including gender, ethnicity, expert group (ecosystem, economy, or public policy), reading comprehension, oral English fluency, talkativeness, good idea nominations, leader nominations, information centrality, and ratings of peer liking. The results indicated that Latino/a children were more likely to produce causal links than African American children, $\beta = 0.48$, $t(3101) = 3.22$, $p = .001$. Students from the economy expert group were more likely to produce causal chains than ecosystem students, $\beta = 0.51$, $t(3101) = 4.49$, $p < .001$ and public policy students, $\beta = 0.59$, $t(3101) = 4.57$, $p < .001$, while these two groups did not differ, $t(3101) = 0.61$, $p = .54$. Students who had faster oral English fluency, $F(1, 3101) = 17.02$, $p < .001$, who were more talkative, $F(1, 3101) = 34.71$, $p < .001$, and who received higher peer liking ratings from classmates, $F(1, 3101) = 21.69$, $p < .001$, showed a higher probability to produce causal links during collaborative discussion.

Next, the number of causal links produced at Lag-1 and Lag-2 were included in the SDA model to investigate whether producing causal links in the previous two turns increased the likelihood of producing causal links in the current turn. The results indicated that the more causal links were produced at Lag-1 and Lag-2, the more causal links were produced in the current turn, $\beta_{Lag\text{-}1} = 0.26$, $F(1, 3051) = 100.71$, $p < .001$ and $\beta_{Lag\text{-}2} = 0.24$, $F(1, 3051) = 75.74$, $p < .001$. This result suggested that once a chain was started, it was likely to continue for three speaking turns.

In Model 3, same sets of social and cognitive characteristics of Lag-1 and Lag-2 speakers were included as explanatory variables to examine whether the previous speakers' social and cognitive characteristics had an impact on the current speaker's multilink causal reasoning. The results indicated that if the previous speaker was a leader in the class or had high information centrality in the classroom network, the current speaker was more likely to produce causal links, $\beta_{leader(Lag\text{-}1)} = 0.10$, $F(1, 3047) = 5.30$, $p = .021$ and $\beta_{centrality(Lag\text{-}1)} = 0.12$, $F(1, 3047) = 3.95$, $p = .047$. This finding suggested that leaders and socially centered students tended to support socially peripheral students to perform multilink causal reasoning during the discussion. Another finding was that if the speaker at Lag-2 had high oral English fluency, it was more likely for the current speaker to produce causal links, $\beta_{English\ (Lag\text{-}2)} = 0.009$, $F(1, 3047) = 4.21$, $p = .040$. This finding suggested that high English proficiency students may support their less fluent peers to express their thoughts in the discussion. The SDA analysis also showed that if the previous speaker had low reading comprehension, it was more likely for the current speaker to extend the chain of reasoning, $\beta_{reading\ (Lag\text{-}1)} = -0.015$, $F(1, 3047) = 6.97$, $p = .008$. This finding suggested that good readers may help poor readers to articulate their thinking during collaborative discussion.

In Model 4, students' extending moves at Lag-1 were entered to explore the influence of peer interaction between previous speakers on the current speaker's multilink causal reasoning. Extending moves at Lag-1 indicated whether the previous speaker supported, refuted, changed the topic, or took a neutral action (e.g., asking a clarification question or manage turn-taking) as a response to the speaker at Lag-2. Although extending moves were made by the Lag-1 speaker, these moves were triggered by the Lag-2 speaker and indicated whether there was a shared understanding between the two speakers. Lin et al. (2015) found that if the previous speaker agreed with the speaker at Lag-2, the current speaker was more likely to extend the previous speaker's thinking. The results of Model 4 corroborated Lin's finding—agreement at Lag-1 significantly predicted the number of causal links in the current turn of speaking, $\beta_{agree} = 0.31$, $F(1, 3044) = 6.83$, $p = .009$, while disagreement at Lag-1 showed a negative effect, $\beta_{disagree} = -0.37$, $F(1, 3044) = 5.10$, $p = .024$. Additionally, change of topic at Lag-1 triggered the production of causal links in the current turn, $\beta_{new\ topic} = 0.43$, $F(1, 3044) = 11.90$, $p < .001$.

Finally, an additional SDA model (Model 5) was created to take repeated and alternating turns into account. In this dataset, repetition of turn-taking occurred when the teacher interrupted a child' talk and the child continued after the teacher's turn. Alternation of turn-taking refers to the situation when the current turn and Lag-2 turn are generated by the same person. A total of 72 repeated turns and 926 alternating turns were observed in the transcripts. The results indicated that multilink causal reasoning was less likely to happen in repeated turns, $\beta_{repeated} = -0.82$, $F(1, 3042) = 5.50$, $p = .019$, or in alternating turns, $\beta_{alternating} = -0.44$, $F(1, 3042) = 16.50$, $p < .001$.

Table 9

Proximal Effects of Previous Speakers' Characteristics and Extending Moves

Effect		Model 1	Model 2	Model 3	Model 4	Model 5
Current speakers' social and cognitive characteristics						
Intercept		-3.69***(0.40)	-3.69***(0.39)	-3.97*** (0.46)	-4.04*** (0.47)	-4.13*** (0.48)
Ethnicity: *Latino/a vs. African American*		0.48** (0.15)	0.49** (0.15)	0.47*** (0.15)	0.48** (0.15)	0.48** (0.15)
Expert group	*Ecosystem vs. Economy*	-0.51*** (0.11)	-0.45***(0.12)	-0.42*** (0.12)	-0.44*** (0.12)	-0.46*** (0.12)
	Economy vs. Public policy	0.59*** (0.13)	0.56***(0.13)	0.56*** (0.13)	0.54*** (0.13)	0.57*** (0.13)
	Ecosystem vs. Public policy	0.08 (0.13)	0.11 (0.13)	0.14 (0.13)	0.11 (0.13)	0.11 (0.13)
Oral English fluency		0.019***(.005)	0.016***(.005)	0.014**(.005)	0.015**(.005)	0.017***(.005)
Talkativeness		0.37***(0.06)	0.40***(0.06)	0.40*** (0.06)	0.40*** (0.06)	0.41*** (0.06)
Peer liking rating		0.26***(0.06)	0.25***(0.06)	0.25*** (0.06)	0.25*** (0.06)	0.25*** (0.06)
Previous speakers' multilink causal reasoning						
Number of causal links (Lag-1)			0.26*** (0.03)	0.27*** (0.03)	0.24*** (0.03)	0.25*** (0.03)
Number of causal links (Lag-2)			0.24*** (0.03)	0.24*** (0.03)	0.23*** (0.03)	0.23*** (0.03)
Previous speakers' social and cognitive characteristics						
Reading comprehension (Lag-1)				-0.015**(0.006)	-0.014*(0.006)	-0.014*(0.006)
Leadership nomination (Lag-1)				0.10* (0.04)	0.09* (0.04)	0.09* (0.04)
Information centrality (Lag-1)				0.12* (0.16)	0.12* (0.16)	0.13* (0.16)
Oral English fluency (Lag-2)				0.009*(0.004)	0.007†(0.004)	0.010*(0.004)
Previous speaker's extending moves						
Support/Agree (Lag-1)					0.31** (0.12)	0.27* (0.12)
Refute/Disagree (Lag-1)					-0.37* (0.16)	-0.37* (0.17)
New topic (Lag-1)					0.43***(0.12)	0.36** (0.12)
Controlling for repeated and alternating turns						
Repeated turn (Lag-1) (-)						-0.82* (0.35)
Alternating turn (Lag-2) (-)						-0.44*** (0.11)
Random effects (Variance of intercept)		0.91** (0.34)	0.55** (0.22)	0.62** (0.25)	0.63**(0.26)	0.58**(0.24)
Fit Statistics (AIC)		3346	3088	3078	3051	3033

Note. Lag–N refers to N turns before the current turn. † p <.10, * p <.05, ** p <.01, *** p <.001

4.4.2 The proximal effects of leadership moves

Leadership moves were identified in 798 speaking turns, including 626 student turns and 172 teacher turns. Approximately twenty percent of the student talk and nearly 90% of the teacher turns were leadership moves. Majority of student-generated leadership moves focused on turn management (32%), argument development (29%) and clarifying and revoicing (24%), while majority of teachers' leadership moves were planning and organizing (35%) and argument development (21%). The average number of leadership moves in each discussion is 33, with a range of 5 to 85. The distribution of leadership moves was presented in Figure 10.

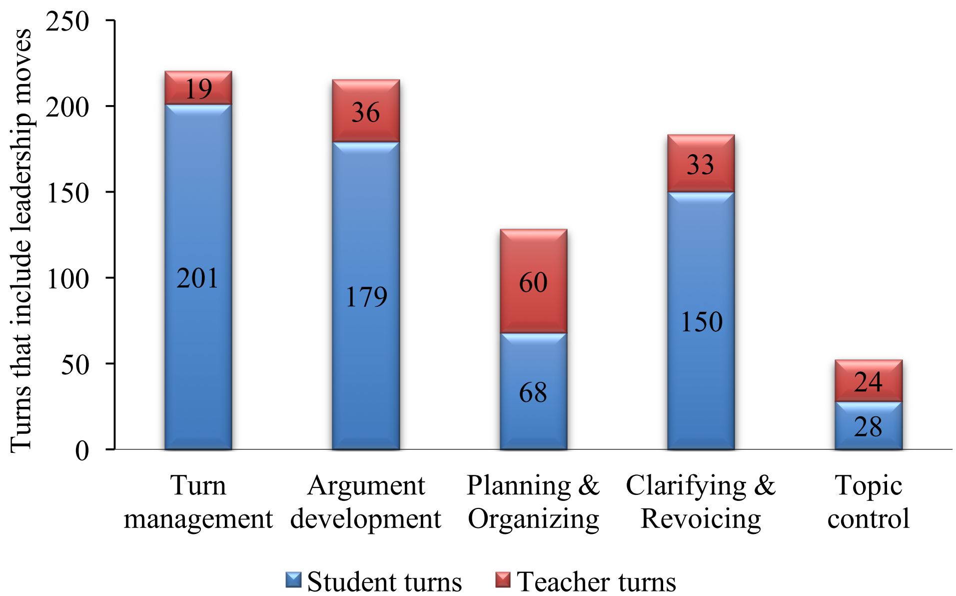

Figure 10. The distribution of five types of leadership move among students and teachers.

A majority (70%) of the students made at least one leadership move during the discussion, while 48 students (30%) did not produce any leadership moves. Leadership moves were widely distributed among group members in 19 groups (nearly 80% of the total). In these groups, on average 87% of the groups members produced at least one leadership move during the discussion (the range is 60% to 100%). Notably, all of the group member produced at least one leadership move during the discussion in eight groups (33% of the total). In the five groups that had low frequency of student-initiated leadership moves, majority (80%) of the leadership moves were generated by the teacher, while in the 19 groups that had high frequency of student-initiated leadership moves, the percentage of teacher-generate leadership moves was only 18%. These findings revealed that student-initiated leadership moves were more likely to appear when students had more autonomy over the management of discussion.

A second statistical discourse analysis (SDA, Chiu, 2008) was performed to investigate the influence of student-generated and teacher-generated leadership moves on turn-by-turn chain construction. Teacher turns were recoded to indicate the leadership function of the speaking turn and its lag position. Teacher turns were first coded as binary lag variables, which indicated the occurrence of an event in previous turns. For example, if a student turn is generated right after a teacher's turn, a binary code (1= a teacher's turn, 0 = not a teacher's turn) is assigned at the turn before the current turn (Lag-1). Likewise, a binary code is assigned to Lag-2 if teacher's talk occurs two turns before the current turn. Next, lag variables were categorized based on its leadership function in the discussion; for example, if a teacher generated a turn management move before the current speaking turn, it is coded as *Lag-1 Turn Management*.

The final SDA analysis included 3132 speaking turns, after recoding all the teacher turns ($N = 193$) and removing unidentifiable student turns ($N = 29$). The dependent variable was the number of causal links produced by a child at the current speaking turn (Lag-0). Two-level SDA models were constructed with speaking turns nested within groups. Intercepts were allowed to vary at the group level. Poisson regression analysis was used because the outcome variable followed a Poisson distribution. To control for individual differences in social and cognitive attributes, the current speaker's ethnicity, expert group (ecosystem, economy, or public policy), basic English proficiency, talkativeness, and ratings of peer liking were entered as Lag-0 variables. Whether or not a leadership move was generated by a student or by a teacher at Lag-1, Lag-2, and Lag-3 were sequentially entered to the SDA model. Non-significant variables were removed from the final SDA model following the rule of parsimony. The results of the final model were presented in Table 10.

Table 10

Proximal Effects of Leadership Moves on Multilink Causal Reasoning

Effect		Estimate (*SD*)
Intercept		-3.70^{***} (0.41)
Current speakers' social and cognitive characteristics		
Ethnicity:	*Latino/a vs. African American*	0.56^{***} (0.15)
	Ecosystem vs. Economy	-0.43^{***} (0.12)
Expert group	*Economy vs. Public policy*	0.57^{***} (0.13)
	Ecosystem vs. Public policy	0.14 (0.13)
Basic English proficiency		0.018^{***} (.005)
Talkativeness		0.39^{***} (0.07)
Peer liking rating		0.25^{***} (0.06)
Student's leadership moves		
Argument development (Lag-1)		-0.64^{*} (0.26)
Clarifying and Revoicing (Lag-1)		-0.67^{*} (0.27)
Turn management (Lag-3)		0.35^{*} (0.16)
Teacher's leadership moves		
Turn management (Lag-1)		1.82^{***} (0.32)
Argument development (Lag-1)		1.44^{***} (0.40)
Topic control (Lag-2)		1.79^{***} (0.29)
Argument development (Lag-2)		$0.80^{\dagger}$ (0.50)
Argument development (Lag-3)		0.91^{**} (0.31)
Random effects		
Variance of intercept		0.99^{**} (0.36)

Note. Lag–N refers to N turns before the current turn. $^{\dagger}p$ <.10, $^{*}p$ <.05, $^{**}p$ <.01, $^{***}p$ <.001

The results indicated that teacher's argument development move had a lasting positive effect on students' chain construction—students were found to produce more causal links right after the teacher prompted for opinions, reasons, or evidence, $\beta = 1.44$, $F(1, 3021) = 12.96$, $p <$ .001, two turns later, $\beta = 0.80$, $F(1, 3020) = 2.95$, $p = .086$, and three turns later, $\beta = 0.91$, F(1,

3021) = 8.97, p = .003. Students were also more likely to produce causal links following a turn management move by the teacher, β= 1.82, F(1, 3021) = 32.73, p < .001, or two turns after the teacher made a topic control move, β= 1.79, F(1, 3021) = 37.60, p < .001.

The effects of students' leadership moves were not as strong as the teachers' leadership moves. Three turns after a group member made a turn management move to invite others to talk, causal links were likely to be observed in that turn, β = 0.35, F(1, 3021) = 4.79, p = .029. However, when the previous speaker made an argument development move, β= -0.64, F(1, 3021) = 6.20, p = .013, or clarifying and revoicing move, β= -0.67, F(1, 3021) = 5.99, p = .014, it was less likely for the current speaker to produce causal links.

4.4.3 The proximal effects of reasoning moves

In open-format, student-managed collaborative discussions, each speaker takes a different perspective on how to develop a unique thread of conversation to move the discussion forward. Among the 3354 full turns of speaking, 1482 turns (44%) were attributed to the justification for killing or not killing wolves, 798 turns (24%) contained leadership moves that contributed to discussion management, 302 turns (9%) focused on the evaluation of evidence, 495 turns (15%) were used to discuss alternate solutions, and 277 turns (8%) were about other topics.

A third statistical discourse analysis (SDA, Chiu, 2008) was performed to investigate the proximal effects of reasoning moves on turn-by-turn chain construction. The final SDA analysis included 3132 speaking turns, after removing all the teacher turns (N = 193) and unidentifiable student turns (N = 29). The dependent variable was the number of causal links produced by a child at the current speaking turn (Lag-0). Two-level SDA models were constructed with speaking turns nested within groups. Intercepts were allowed to vary at the group level. Poisson regression analysis was used because the outcome variable followed a Poisson distribution. To control for individual differences in social and cognitive attributes, the current speaker's ethnicity, expert group (ecosystem, economy, or public policy), basic English proficiency, talkativeness, and ratings of peer liking were entered as Lag-0 variables. Whether or not a speaker generated one of the four types of reasoning moves (i.e., positioning and justifying for wolf hunting, evaluating the credibility of evidence, proposing alternate solutions, and discussing other topics) at Lag-1, Lag-2, and Lag-3 were sequentially entered to the SDA model. Non-significant variables were removed from the final SDA model following the rule of parsimony. The results of the final model were presented in Table 11.

Table 11

Proximal Effects of Reasoning Moves on Multilink Causal Reasoning

Effect		**Estimate (*SD*)**
Intercept		-3.38^{***} (0.39)
Current speakers' social and cognitive characteristics		
Ethnicity: *Latino/a vs. African American*		048^{**} (0.15)
	Ecosystem vs. Economy	-0.44^{***} (0.12)
Expert group	*Economy vs. Public policy*	0.51^{***} (0.13)
	Ecosystem vs. Public policy	0.08 (0.13)
Basic English proficiency		0.017^{***} (.005)
Talkativeness		0.38^{***} (0.06)
Peer liking rating		0.25^{***} (0.06)
Lag-1 predictors		
Positioning and justifying for wolf hunting (+)		0.49^{***} (0.11)
Proposing alternate solutions (-)		-1.19^{***} (0.33)
Lag-2 predictors		
Proposing alternate solutions (-)		-0.80^{**} (0.29)
Evaluating the credibility of evidence (-)		-1.05^{*} (0.43)
Lag-3 predictors		
Evaluating the credibility of evidence (-)		-1.36^{**} (0.46)
Discussing other topics (-)		-1.66^{***} (0.37)
Random effects		
Variance of intercept		0.53^{**}(0.21)

Note. Lag–N refers to N turns before the current turn. "+" = positive effect, "-" = negative effect. Significance level: $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

The results of the SDA model showed that if the previous speaker took a position on the wolf killing question and provided a justification for the position, the current speaker was more likely to continue the line of reasoning and produce causal links, $\beta = 0.49$, $F(1, 3023) = 21.99$, $p < .001$. However, if the previous speaker proposed an alternate solution to the wolf problem, the current speaker was less likely to produce causal links, $\beta = -1.19$, $F(1, 3023) = 12.91$, $p < .001$. The negative effect of proposing an alternate solution was also significant at Lag-2, $\beta = -0.80$, $F(1, 3023) = 7.61$, $p = .006$, meaning that the two speaking turns following an alternate solution were less likely to contain causal links. Multilink causal reasoning was also less likely to appear when previous speakers questioned the wolf threat, as shown by the negative effect of evaluating the credibility of evidence at Lag-2, $\beta = -1.05$, $F(1, 3023) = 6.04$, $p = .014$, and at Lag-3, $\beta = -1.36$, $F(1, 3023) = 8.83$, $p = .003$. Additionally, students were less likely to produce causal links if a group member went off-topic or started to share wolf-related knowledge that strays away from the central question three turns ago, $\beta = -1.66$, $F(1, 3023) = 20.49$, $p < .001$.

The findings from the SDA analysis suggest that multilink causal chains were more likely to appear when students justified their position on wolf hunting based on text information, personal experiences, or their understanding of moral principles. However, multilink causal chains were less likely to appear while they were evaluating the evidence of wolf's threat, proposing alternate solutions, or discussing other topics.

4.5 A Causal Analysis of Peer Modeling and Individual Outcome

Children's multilink causal reasoning was also investigated in a post-intervention written task in which children expressed their opinions and decided on whether the townspeople should hire professional hunters to kill the wolves. Seven multilink causal reasoning models were tracked in children's policy decision letters. The percentage of agreement between two coders was 93% for the coding of multilink causal reasoning in decision letters (Cohen's K = 0.88). The results showed that a total number of 89 causal chain were produced by 66 students (43%) in the policy decision letters. The length of chains in the letters ranged from three links to eight links. The total number of causal chains, the length of causal chains and the total number of causal links were calculated for each individual. The present study took four steps to understand how children's production of causal chains in collaborative discussion affect their performance in the independent written task. All the analyses were conducted in Stata 14.

First, a multiple regression analysis was conducted to examine what kinds of individual-level social and cognitive characteristics affected children's production of multilink causal chains in the decision letter. The number of causal links produced by each individual was treated as the outcome variable. Since the outcome variable was count data that followed a Poisson distribution, a Poisson regression model was constructed with gender, ethnicity, expert groups, and whether or not a student was enrolled in IEP program as fixed effects. Individual-level reading comprehension, oral English fluency, talkativeness, leader nominations, information centrality, and peer liking ratings were entered as covariates. The final model (non-significant predictors were dropped) was presented in Figure 11.

The Poisson regression analysis indicated a significant ethnicity difference—Latino/a children produced more causal links in their decision letters than African American children, $\chi^2(N = 154) = 59.30$, $p < .001$. Students in the economy group produced more causal links than students in the public policy group, $\chi^2(N=154) = 6.58$, $p = .010$. Talkative children, $\beta_{\text{talkative}}=0.25$, $\chi^2(N=154) =15.98$, $p < .001$, children who speak good English, $\beta_{\text{English}} = 0.40$, $\chi^2(N=154) =35.81$, $p < .001$, and children who have high reading ability, $\beta_{\text{reading}}= 0.18$, $\chi^2 (N=154) = 10.89$, $p = .001$, produced more causal links in the wolf letter, as compared to quiet children, poor readers, and children with low English fluency. Surprisingly, information centrality had a negative effect on children's chain production in the wolf letter, $\beta_{centrality} = -0.28$, $\chi^2(N = 154) = 18.59$, $p < .001$.

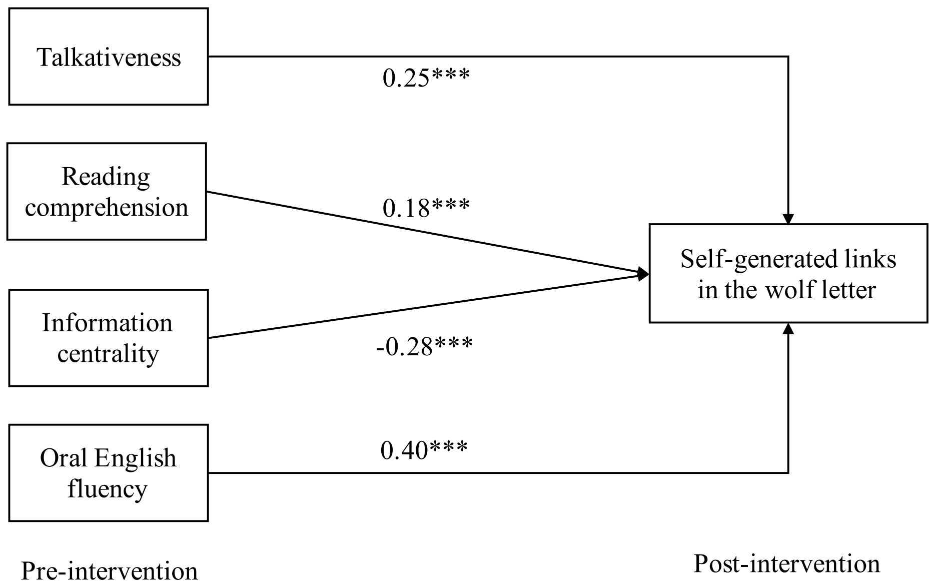

Figure 11. The Poisson model for analyzing the effect of individual characteristics on children's chain production in the wolf letter. * $p < .05$, ** $p < .01$, *** $p < .001$.

In the second analysis, number of self-initiated causal links during collaborative discussion was entered to the Poisson model to examine whether building causal chains during group discussion improved children's ability to generate causal chains in the decision letter. Number of initiated links was used because it indicates spontaneous production of causal links rather than prompted production. The total number of causal links produced by each child in the wolf letter was used as the outcome variable. Talkativeness, reading comprehension, oral English fluency, and information centrality were used as the explanatory variables. Ethnicity and expert group were entered as fixed effects. Non-significant predictors such as gender, whether or not students enrolled in IEP, leader nomination, and peer liking were not included in the model. Since children's social and cognitive characteristics also affected their performance in group discussion, a path model (see Figure 12) was created with children's social and cognitive attributes being the predictors for chain production in both the discussion and the wolf letter.

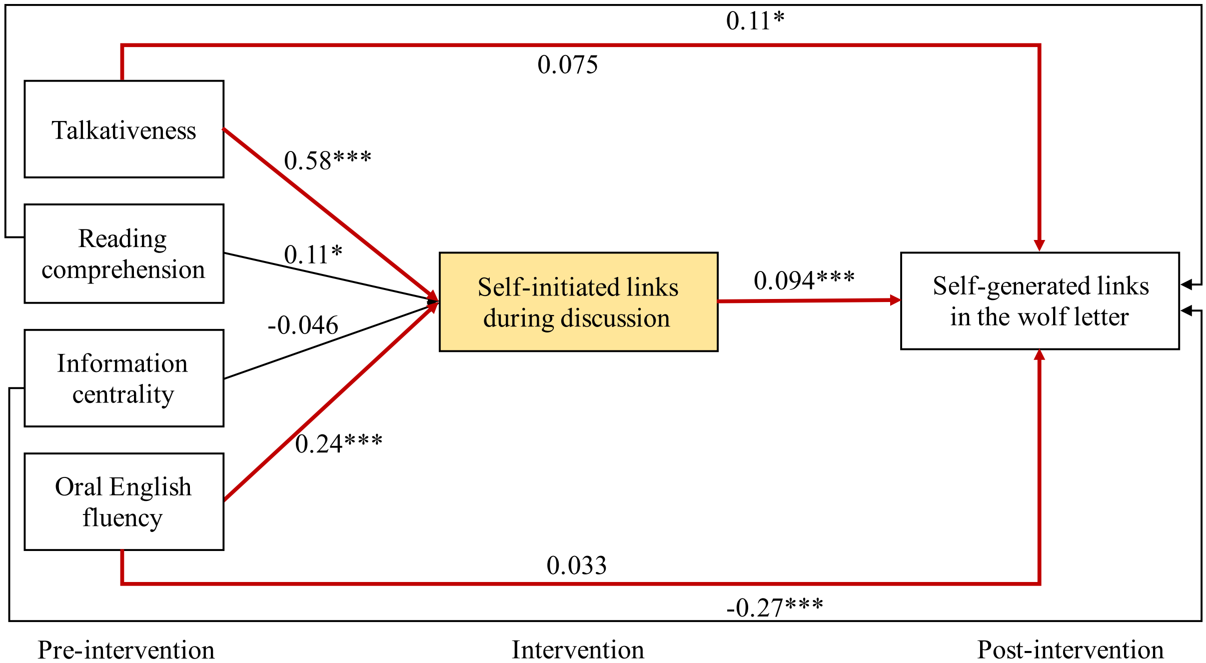

Figure 12. The path model for analyzing the effect of self-initiated links during discussion on children's chain production in the wolf letter. * $p < .05$, ** $p < .05$, *** $p < .001$.

The path model indicated that children who initiated more causal links during group discussion also generated more causal links in the wolf letter, $\beta = 0.094$, $Z = 7.15$, $p < .001$. After incorporating self-initiated links, the effect of talkativeness on chain production in the wolf letter was no longer significant, $\beta = 0.075$, $Z = 1.21$, $p = .23$, as well as the effect of oral English fluency on the outcome measure, $\beta = 0.033$, $Z = 0.56$, $p = .58$. Instead, there was a significant effect of talkativeness, $\beta = 0.58$, $Z = 9.68$, $p < .001$, and a significant effect of oral English fluency, $\beta = 0.24$, $Z = 4.17$, $p < .001$, on the production of causal links during group discussion. Reading comprehension was a significant predictor for chain production in the discussion, $\beta = 0.11$, $Z = 2.09$, $p = .037$, as well as for chain production in the letter, $\beta = 0.11$, $Z = 2.03$, $p = .042$. Information centrality had a negative effect on children's chain production in the letter, $\beta = -0.27$, $Z = -4.41$, $p < .001$, but showed no effect on chain production in the discussion, $\beta = -0.046$, $Z = -0.68$, $p = .50$. The results of the path model suggested that talkative children and children with high English fluency had more opportunity to explicitly generate multilink causal chains during collaborative discussion, and in turn performed better in the following written task, as compared to quiet children or children with low English fluency.

In the third analysis, number of peer-initiated causal links during collaborative discussion was entered to the path model to examine the effect of peer performance on children's multilink causal reasoning in the wolf letter. Peer-initiated links was calculated as the sum of causal links initiated by all the group members except the participant during the discussion. The path model was presented in Figure 13 (non-significant predictors were not shown in the figure).

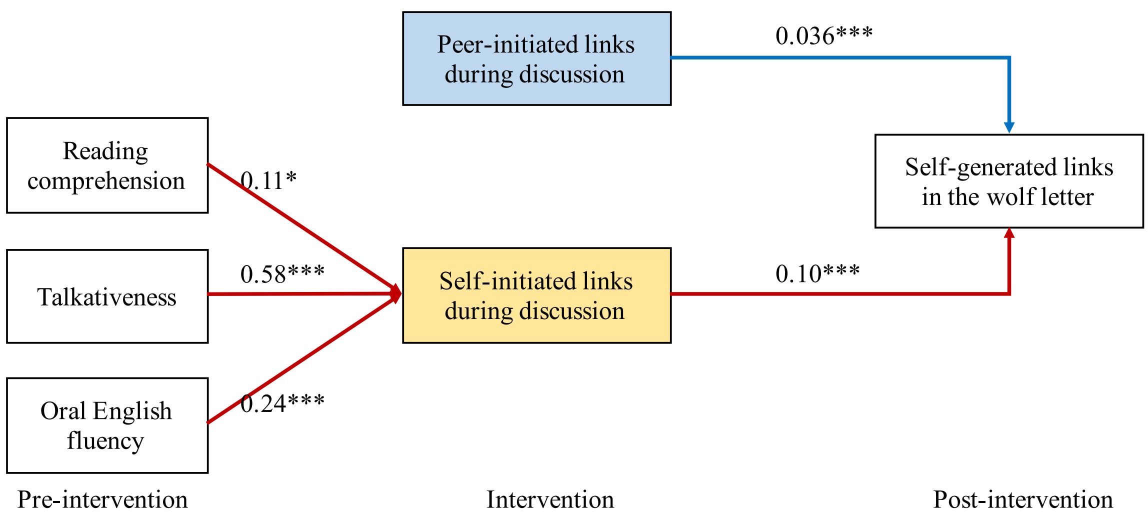

Figure 13. The path model for analyzing the effect of peer-initiated links during discussion on children's chain production in the wolf letter. * $p < .05$, ** $p < .01$, *** $p < .001$.

Peer-initiated links during group discussion showed a significant positive effect on children's multilink causal reasoning in the wolf letter, $\beta = 0.036$, $Z = 5.17$, $p < .001$. The effect of self-initiated links remained significant after incorporating number of peer-initiated links, $\beta = 0.10$, $Z = 7.36$, $p < .001$. Talkativeness, $\beta = 0.58$, $Z = 9.68$, $p < .001$, reading comprehension, $\beta = 0.11$, $Z = 2.09$, $p = .037$, and oral English fluency, $\beta = 0.24$, $Z = 4.17$, $p < .001$, remained as significant predictors of chain production during the discussion.

Since both self-initiated links and peer links significantly predicted the production of multilink causal chain in the wolf letter, an interaction term was added to the path model to examine if the effect of self-initiated links varies at different level of peer-initiated links. Figure 14 showed the path model that included the interaction of self-initiated links and peer-initiated links. As shown in Figure 15, there was a significant interaction of self-initiated links and peer-initiated links, $\beta = 0.006$, $Z = 2.28$, $p = .023$. After including the interaction term, the effect of self-initiated links became insignificant, $\beta = 0.039$, $Z = 1.28$, $p = .20$. However, the effect of peer-initiated links remained significant, $\beta = 0.025$, $Z = 2.98$, $p = .003$.

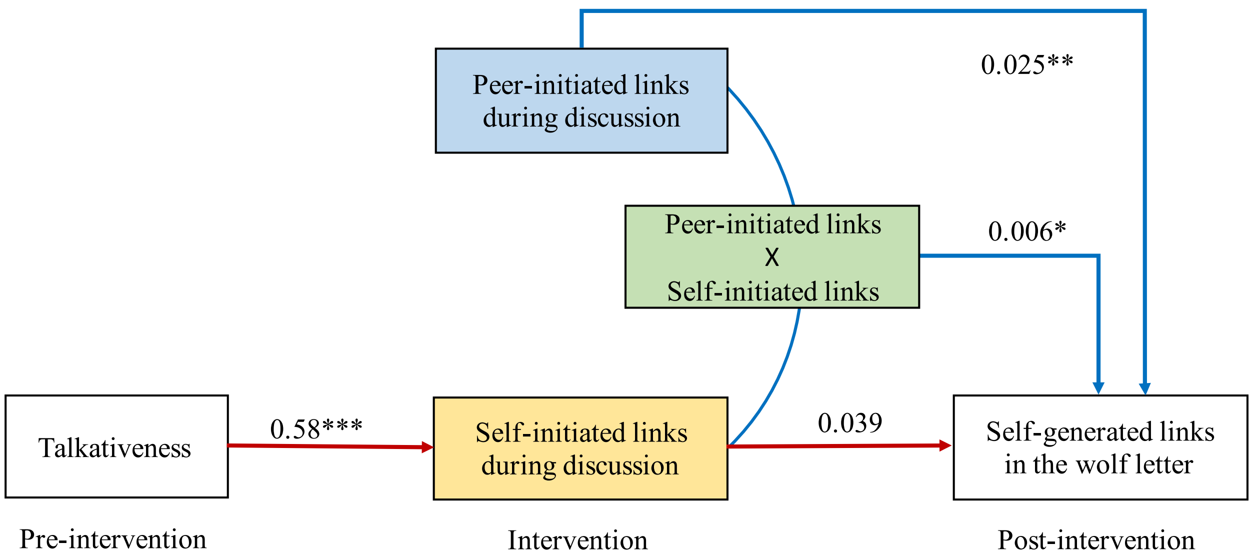

Figure 14. The path model for analyzing the interaction effect of self-initiated links and peer-initiated links on children's chain production in the wolf letter. * p <.05, ** p <.01, *** p < .001.

This analysis suggests that both self-initiated links and peer-initiated links during collaborative discussion influenced children's ability to generate causal chains in the wolf letter; however, number of peer-initiated links moderated the effect of self-initiated links—the effect of self-initiated links was not significant when group members did not produce any causal chains, but as the number of peer-initiated links increased, there was a significant growth of the effect of self-initiated links on chain production in the letter.

Finally, **a fourth analysis** was conducted to examine if the effect of peer-initiated links differed between talkative children and quiet children. As suggested by previous analyses, talkative children were more likely to openly express their ideas during collaborative discussion and therefore had more opportunities to explicitly practice multilink causal reasoning during the discussion. Quiet children, in contrast, spent most of the discussion time listening to their peers, so not much use of multilink causal reasoning was observed among these children during the discussion. However, I wonder if quiet children can benefit from hearing peers talking about adding or removing links in causal chains, as compared to sitting with a group of children who produced little or no chains. My hypothesis is that talkativeness has a stronger effect on chain production in the letter if number of peer-initiated links is low; however, talkativeness is less influential if number of peer-initiated links is high.

Following the analytic method in Hayes (2013), the interaction of talkativeness and peer-initiated links was added to the previous path model. In the new path model (see Figure 15), peer-initiated links moderated both the effect of self-initiated links and the effect of talkativeness on the number of causal links produced by each child in the wolf letter. When group members did not produce any causal chains during the discussion (i.e., number of peer-initiated links equals to zero), the effect of self-initiated links on children' chain production in the wolf letter was not significant, $\beta = 0.006$, $Z = 0.18$, $p = .85$. In this situation, the more talkative a child was, the more likely that this child was able to produce causal chains in the wolf letter, $\beta = 0.27$, $Z = 2.22$, $p = .027$. However, as the number of peer links increased, the effect of self-initiated links increased, $\beta = 0.009$, $Z = 3.02$, $p = .003$, and the influence of talkativeness decreased, $\beta = -0.017$, $Z = -2.11$, $p = .035$. In other words, when a quiet child did not hear many causal chains from the sharing of group mates during the discussion, it was less likely for this child to produce causal links in the letter. However, if this quiet child sat in a group in which group members generated many causal links, it was more likely for this child to produce causal links in the wolf letter.

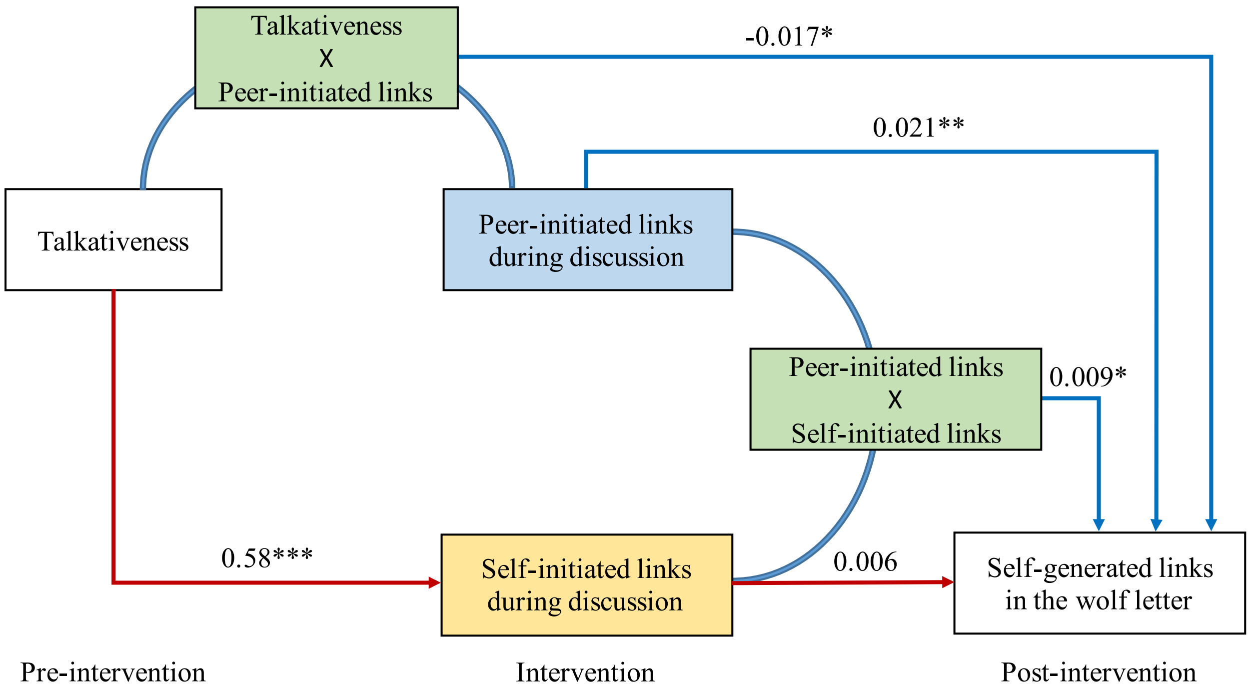

Figure 15. The path model for analyzing the interaction effect of peer-initiated links and talkativeness on children's chain production in the wolf letter. * $p < .05$, ** $p < .01$, *** $p < .001$.

Taken together, children who produced more causal links during the discussion also produced more causal links in the letter. However, for children who did not produce many causal chains themselves, if their peers produced many causal links during the discussion, it is more likely for them to internalize others' thinking into their own ideas, especially for quiet children.

Previous findings have shown that some groups spent more time on proposing alternate solutions, evaluating evidence of wolf threat, and discussing other topics, which may affect the number of causal links that the group can produce. So, I reran the causal analyses after controlling for the number of speaking turns that each group used to discuss topics that do not directly address the wolf question. The adjusted analyses showed similar patterns of results but stronger statistical significance and larger effects due to more sensitivity.

CHAPTER 5

DISCUSSION

The major contribution of this dissertation is to document the psychological and social reality of the concept of multilink causal reasoning. The study illuminates many aspects of the process of multilink reasoning, how the process works, and the roles played in the process by different kinds of children, which I will address in the remainder of this chapter. But before I plunge into the task of summarizing and interpreting the findings, I want to say that it is a dream come true to be able to complete a pioneering study of an important phenomenon. The study that has yielded a number of findings that, as far as I know, are original to the field.

To recap, this study characterizes children's discussion groups that produce many multilink causal chains, the kinds of children who initiate the chains, the characteristics of children who add or repeat links, and the features of dialogue that support the construction of causal chains. The study also examines the proximal effect of dynamic, reflexive social interaction on children's chain construction during collaborative discussion, tracks the temporal development of multilink causal chains over the course of a discussion, and investigates the application of multilink causal reasoning in a written task following the collaborative discussion.

In the remainder of this chapter, I will summarize the major findings and try to explain how multilink causal reasoning develops as a social process. To further understand this social process, I will describe the effect of moment-by-moment peer interaction on children's application of multilink causal reasoning in an individual written task, and I will reflect on the intrapersonal and interpersonal factors that affect the social construction of multilink causal chains. The present chapter also discusses the educational implications of this study, considers the limitation of the analyses, and proposes future directions for research.

5.1 Major Findings on the Development of Multilink Causal Reasoning

The first major finding of this study is that students spontaneously generated multilink causal chains through interacting with peers during collaborative discussions. The construction of multilink causal chains was distributed among group members and was spread across speaking turns. Recurrent patterns of multilink causal reasoning were identified in a majority (92%) of the collaborative discussions. Sixty percent of the children were found to participate in the production of multilink causal chains. Forty percent of the causal chains were co-constructed by several children; one child initiated the chain and the other children added links to the chain or elaborated certain links. The turn-by-turn analysis (see Section 4.4.1) of chain construction showed that once a causal chain was started, it was likely to continue for three speaking turns.

The spread of multilink causal reasoning across groups of children substantiated the *snowball phenomenon,* or peer modeling effect. The snowball phenomenon happens when children adopt the way other people think and talk, once they find these strategies to be effective, to enhance their own talk during group discussions. Once students have picked up a new reasoning or rhetorical strategy by imitating their peers, they are often able to use the strategy on their own. The snowball phenomenon was first reported in Anderson et al. (2001), in which each of several different argument stratagems was found to spread across groups of children and occur with increasing frequency after its first appearance. The snowball phenomenon has been observed in the use of argument stratagems during collaborative discussions by Chinese and Korean children (Dong et al., 2008) and in the uptake and use of several kinds of relational thinking by American students in collaborative groups (Lin et al, 2012; Lin, Jadallah et al., 2015). Besides replicating the peer modeling effect shown in previous studies, the noteworthy contribution of this study is that it substantiates the snowball phenomenon in the spread of hierarchical reasoning patterns rather than the spread of a specific reasoning strategy. The significance of this finding is that it suggests that social propagation may be a general mechanism in the development of reasoning.

The second major finding is that the rate of generating multilink causal chains increased over the course a collaborative discussion. As shown by the temporal analysis (see Section 4.3), each successive chain took less and less time to generate, starting with the second chain. The snowballing of multilink causal reasoning at an increasing rate suggests that conceptual relationships, as well as the words to express these relationships, may become more accessible and widely available to group members over the course of a discussion. The results of the event history analysis (see Section 4.3) showed that long causal chains were more likely to appear in a late stage of discussion and the number of speaking turns needed to produce a causal chain decreased as the discussion progressed. A possible reason for this set of findings is that concepts and relationships were not widely available to group members in earlier stages of the discussion and therefore students had to take more speaking turns to negotiate whether certain links should be added to the chain. As students made public more concepts and relationships, they were more capable of integrating ideas from multiple sources into a coherent causal chain. The decrease of speaking turns for generating a causal chain further suggests that the representation of lexical and syntactic information becomes more integrated and coordinated in students' minds over time.

According to Kintsch's (1988, 1998, 2005) prominent Construction-Integration theory of comprehension, texts have three levels of representation, i.e., surface forms, text-based inferences, and situation models. Surface forms refer to the semantic meaning of the text. Text-based inferences refer to the interpretive meaning of the text, which is often inferred from the relationship between elements of surface forms. The highest level is the situation model. The situation model goes beyond the literal and interpretative meaning of the text and integrates all the surface forms and text-based inferences into a coherent mental representation (Perring & Kintsch, 1985; Zwaan & Radvansky, 1998). Multilink causal chains, in this sense, are situation models that integrate all the surface forms and text-based inferences.

Kintsch (1988) proposed a construction-integration model to explain how learners construct situation models in the course of text comprehension—the first phase involves the construction of a knowledge network that contains all associative information with or without regard to the discourse context, and the second phase carries out an integration process in which irrelevant information is filtered out and the most closely related information is integrated into a coherent mental representation. Similar ideas can be applied to the construction of multilink causal chains. Simple relations are created based on surface forms and causal-relations are the integration of text-based inferences. The construction of multilink causal chains, therefore, is achieved through a construction process in which students come up with words, inferences, or elaborations that are related to the main topic, and then the most relevant ideas are linked together to form a coherent causal chain during the integration process.

In the early stage of a discussion, information is likely to be presented in either surface forms (e.g., "wolves eat elk") or text-based inferences (e.g., "there will be fewer elk"). Chains produced later in the discussion are comprised mostly of links and of two- and three-link assemblies that were constituents of chains generated earlier in the discussion. For example, core propositions such as "Wolves eat elk" and "Wolves kill sheep" are expressed several times in most discussions and typically are constituents of more than one chain. Later restatements repeat or elaborate these core relationships, for example, "Wolves eat elk and deer" and "Wolves kill sheep but not very many of them." Statement and restatement of conceptual relationships makes these relationships accessible to all group members and increases the efficiency with which groups can construct causal chains. This resulted in an observed decrease in time between successive pairs of chains.

Once a situation model in the form of a multilink causal chain is constructed, it usually lasts longer in memory than surface forms or text-based inferences (Kintsch, Welsch, Schmalhofer, & Zimny, 1990). Therefore, short chains can be quickly retrieved from long-term memory when cues (e.g., words such as “wolf” or “elk”) or structures (i.e., relationships such as “wolves eat elk”) are activated in the short-term memory (Ericsson & Kintsch, 1995). Long chains tended to contain the most frequently mentioned causal links, which were used as a basis for forming new links, since students had reached a consensus on the structure.

The third major finding of this study is that students were able to transfer what they had learned about the construction of multilink causal chains in collaborative groups to an independently written essay. The causal analysis indicated that both self-initiated links and peer-initiated links during collaborative interaction boosted children’s chain production in the essay. For children who generated many causal links during group discussion, having the opportunity to hear many peer-initiated links was less important to their use of multilink causal reasoning in the essay. However, for children who did *not* produce many causal links during group discussion, because they were quiet and seldom expressed their opinions, when other members of their group produced many causal links, they were able to internalize what they had hear from their peers and perform multilink causal reasoning in their own writing.

The ability to generate multilink causal chains in independent writing after participating in collaborative discussions indicated that students had developed a general tendency to connect ideas and integrate these ideas into coherent mental models. There are several probable reasons that collaborative discussion fosters children's multilink causal reasoning. A major reason is that the interactive learning environment provides ample opportunity for students to extend each other's talk. Galton, Hargreaves, and Pell (2009) found that students engaged in extensive cognitive discourse during collaborative group work, as indicated by asking more questions, making more suggestions, and providing more explanations of a problem. Morris et al. (forthcoming) found that students who participated in collaborative discussions used more coordinating conjunctions – including *because, so, but, if, then* – in their talk as compared to students who received direct instruction. It is the daily experience over a six-week period of extending other people's talk that leads to the disposition to connect ideas.

The second reason that collaborative discussion fosters multilink causal reasoning is that collaborative group work enables children to share knowledge. Long chains, in particular, cannot be constructed without knowledge sharing. Chains with five or more links were three times more likely to be constructed by a group rather than by an individual. And long causal chains were three times more likely to take two or more speaking turns to generate. The reason that long causal chains take more children and turns to construct is that these chains are usually built on cross-domain knowledge and thus require information from multiple sources. For example, the ecosystem-economy model was built on the interdependent relationship between consumers and producers (ecosystem concepts) and the supply-demand relationship between different types of businesses such as restaurants and wholesalers (economy concepts). During collaborative discussions, students who had become 'experts' in different knowledge domains had the chance to share their expertise with group members and provide their specialized perspective on the wolf question based on what they had learned. Through knowledge sharing, students were able to integrate information from multiple domains of knowledge to explain their views on whether or not to eliminate the wolf pack.

The third reason that collaborative discussion fosters multilink causal reasoning is that collaborative group work provides children sufficient opportunities to elaborate their ideas and explain them to each other. In collaborative discussions, a speaker is often asked to clarify or elaborate. It is expected that each speaker will justify his or her position and defend it when group members challenge it. Webb (2013) documented that students select the most relevant and accurate information to present while working in a group, and provide explanations to others when at first their arguments provoke misunderstandings or lead to misconceptions. Students monitor their own thinking and use other's feedback to create new explanations to resolve cognitive conflicts in order to maintain a shared understanding among group members (Webb, 2013). Renkl, Stark, Gruber, and Mandl (1998) reported that eliciting explanations facilitates knowledge transfer among learners because it helps learners to revisit their thinking and revise their mental models. By explaining an idea to oneself and to others, students develop a better understanding of concepts and relationships, and as a result, acquire the ability to transfer knowledge to a different context.

The fourth probable reason that collaborative discussion fosters multilink causal reasoning is that collaborative activities create communicative space for children that enables them to improve their language proficiency, which enables them to express complex ideas such as causal chains. By talking with peers, children with low language proficiency have the opportunity to learn to speak in more complete, precise, and informative ways. Swain, Brooks, and Tocalli-Beller (2002) and Zhang, Anderson, and Nguyen-Jahiel (2013) found that peer-peer dialogue improved second language learners' proficiency in reading, writing, speaking, and listening. Mackey and Goo (2012) stated that negotiated interactions between language learners and native speakers facilitated language learners' comprehension and vocabulary learning, pushed them to use relevant resources to produce appropriate language output, and provoked them to adjust their language based on academic linguistic forms and conversation structure. In another study, Fernández Dobao (2014) examined the effect of peer interaction on Spanish-speaking English language learners' vocabulary acquisition and found that students who worked in a group of four produced more language-related episodes in which they negotiate the word meanings or the use of words, than students who worked in pairs. By working with more people, learners had access to more knowledge and resources and receive in-time assistance from group members.

Overall, children who participate in collaborative discussions have the opportunity to exchange information with each other and socialize their thinking in an open environment. By working with others, students are hypothesized to form a systematic view that things and events are related and develop a general tendency to search for the antecedents and consequences of events, which therefore increases the possibility of knowledge transfer. However, it is important to emphasize that collaborative group work is more than asking a group of children to talk with each other. In this dataset, there were two groups that did not produce any multilink causal chains. The talk that contributed to the production of multilink causal chains comprised only 7% of the discussions, which suggests that multilink causal reasoning is still a challenging skill for children, especially perhaps for these children who live in underserved communities and have limited access to educational resources.

5.2 The Momentary Social Process in which Multilink Causal Reasoning Emerges

Considered next is the social process in which multilink causal reasoning develops, with the goal of understanding how multilink causal reasoning is nurtured through moment-by-moment social interaction. Many factors influenced children's multilink causal reasoning during collaborative discussions in this study. The intrapersonal factors such as oral English fluency, talkativeness, and reading comprehension affected how many speaking turns a child generated during a group discussion and in turn changed the likelihood of producing a causal chain. And interpersonal factors such as peer assistance or teacher's prompts had a proximal effect on a child's thinking, which increased or decreased the possibility of a child's extending a causal chain during a given speaking turn. Intrapersonal factors interacted with interpersonal factors to affect a child's ability to participate in the construction of multilink causal chains.

During collaborative discussion, cognitively and socially advanced children added more links to causal chains than other children. The analysis of individual-level characteristics showed that talkative children, children with high oral English fluency, and popular children produced more causal chains than quiet children, children with low oral English fluency, and less popular children. And the analysis of group-level characteristics showed that group average reading comprehension had positive correlations with the number of causal chains produced by a group. The reason that cognitively and socially advanced children produced more causal chains is that they had more speaking turns to express their opinions in the discussion and their ideas were more likely to be supported by group members, as compared to their counterparts.

Socially and cognitively advanced children also played a more active role during the construction of causal chains, whereas socially and cognitively less advanced children participated by responding to the requests or invitations of others. The analysis of features of multilink causal chains showed that children with high oral English fluency were more likely to initiate a causal chain or add to a causal chain rather than repeating links in a causal chain as compared to children with low oral English fluency. Information centrality, an indicator of a child's social status in the class, showed a negative relationship with the production of added links and repeated links, which indicates that, whereas high status children initiate chains, low status children are more likely to add to chains or repeat links in chains.

Another intrapersonal factor is the knowledge domain in which children developed expertise. Students were formed into different groups to study one of the three knowledge domains related to the wolves, namely the ecosystem, the economy, or public policy. Individual-level analysis showed that economy students produced the most causal chains because they were more capable of producing bridging links. A bridging link is a link that connects two knowledge domains, a critical component of long, cross-domain causal chains. For example, the bridging link in the ecosystem-economy model (see Figure 2) is that a decrease in tree population will reduce timber production and in turn affect the timber business. The majority of bridging links were built on concepts in the economy domain because these links describe how killing animals (wolf or wolf's prey) affect various types of businesses such as ranching or hunting. Fewer long chains were produced by ecosystem students and public policy students because economy concepts were required to make cross-domain connections.

Besides the aforementioned intrapersonal factors, interpersonal factors also had an impact on children's multilink causal reasoning, and furthermore, several interpersonal factors compensated to some extent for counterproductive intrapersonal factors. As emphasized earlier, quiet children were able to transfer what they had learned from their peers to the independent essay despite talking much less during the group discussion. The turn-by-turn analysis of chain construction below identifies several patterns of interaction of intrapersonal and interpersonal factors.

First, this study identified several characteristics of the previous speaker that were positively related to the extension of a causal chain in the subsequent turn. If the previous speaker had high leadership nomination or high social centeredness in the class, it was more likely for the subsequent speaker to produce causal links. This finding implies that leaders and socially centered children provided assistance and support for quiet and shy children during collaborative discussion. Lin, Anderson et al. (2015) found that students' social status in the classroom mediated the effects of collaborative discussions on relational thinking. Miller et al. (2009) analyzed quiet children's participation in collaborative discussions and found that leaders who emerged in the discussion group successfully encouraged quiet students to talk more. Thus, it is likely that in the current study, quiet children gained opportunities to speak and to enact thinking and reasoning with social support from talkative students and emergent leaders.

However, an interpersonal factor that showed a negative influence on chain production was the previous speaker's reading comprehension. Statistical discourse analysis showed that if the previous speaker had lower reading comprehension, it was more likely for the subsequent speaker to produce causal links. A possible explanation is that poor readers have difficulty articulating their thinking; therefore, group members are likely to take the following speaking turn to explain concepts and elaborate the relationships among them to help the poor reader understand.

Second, this study found that the way the present speaker responded to the previous speaker's ideas or actions affected the subsequent speaker's contribution. Agreement and support among consecutive speakers extended the chain of reasoning while disagreement or refutation stopped the chain of reasoning. Agreement between two speakers acts to encourage more talk along the present line of reasoning. When a person's idea is approved by another person, the idea is validated, receives more attention from group members, and is more likely to be further discussed. Agreement among group members also indicates shared understanding of how concepts or events are related. In Orsolini and Pontecorvo's (1992) study of children's talk in classroom discussions, they found that peer continuations, that is, extension of talk by the subsequent speaker, were more likely to foster continuation of talk beyond that speaker's turn; however, opposition or disputes were more likely to stimulate counter-opposition and rebuttal. Lin, Anderson, et al. (2015) also found that agreement between two previous speakers increased the likelihood of confirmation during the current speaking turn while disagreement between two previous speakers triggered refutation. When disagreement arises, group members have to stop the chain to resolve the disputed idea before they can continue the line of reasoning.

However, the chain of reasoning may be resumed after the group takes several turns to resolve disagreement. Orsolini and Pontecorvo's (1992) performed a turn-by-turn analysis of children's interactions during a classroom discussion and suggested a delayed effect of disagreement on the extension of talk. They claimed that

> "in classroom discussions, children's disputes are embedded discourses. They are prefaced by a teacher's request for explanation or evaluation and are inserted into a wider activity in which the means-end structure of actions has been foregrounded both by a problem-solving task and by the interpretation of the task as being open to different possible solutions" (Orsolini & Pontecorvo, 1992, p. 133).

Following Orsolini and Pontecorvo's (1992) idea, disagreement may temporarily stop the chain of reasoning, but have a delayed effect of prompting students to revisit their thinking in order to resolve the cognitive conflict. Under certain circumstances, disagreement may stimulate the production of new ideas and advance the development of multilink causal chains. However, there are still methodological challenges to demonstrating whether or not resolving disagreement induces later chaining in the discussion, because the point at which a chain will be resumed is difficult to identify. Further analysis is needed to explore whether there are positive delayed effects of disagreement.

Third, this study found that students' and teachers' leadership moves at the previous speaking turn affected the present speaker's production of causal links. The group-level analysis showed that if students in a group had higher average leadership nomination, the group produced more types of causal chains, which suggests that emergent leaders may have prompted other group members to contribute new ideas to the discussion and consequently elicited the production of new causal chains. Groups with higher average leadership nomination also tended to produce more co-constructed chains. This finding suggests that emergent leaders may balance the amount of talk produced by each member of the group. If there were no emergent leaders to negotiate opportunities for shy and quiet children to speak, the discussion may be dominated by talkative children, resulting in the production of more single-person causal chains.

Turn-by-turn analysis revealed that if a student invited a group member to speak, usually a child who had not yet shared his or her opinion, a causal link was likely to be generated within three speaking turns. Teachers' redirection of the topic and management of turn-taking also showed positive effects on student's use of multilink causal reasoning. Analysis of extending moves showed that students' or teacher's redirection of the topic functioned as a resetting mechanism for chaining. When a child or a teacher brought up a new topic, the new idea was likely to spark the production of related ideas (Nijstad, Stroebe, & Lodewijkx, 2002) and thus start a new chain of reasoning. Inviting a group member to speak was likely to result in the initiation of a new topic, which redirected the line of thinking and increased the chances of a new multilink causal chain.

However, the analysis of leadership moves also showed that if a student clarified or restated an argument (revoicing and clarifying move), or asked the previous speaker why and how questions (argument development move), it was less likely for the following speaker to produce a causal link to extend this line of thinking. There are two possible reasons for the negative effect of students' argument development moves and clarifying and revoicing moves on the following speaker's production of causal links. One possible reason is that extending a causal chain depends on the subsequent speaker's response to a leadership move. If the subsequent speaker accepts the leadership move, he or she is likely to provide an elaborated explanation or clarification. However, if the subsequent speaker rejects the leadership move, the previous line of thinking is likely to be terminated. The other possible reason is that the request for explanation or clarification from peers may be regarded as a challenge, rather than an encouragement to extend the line of reasoning. Thus, the subsequent speaker may take the turn to defend his or her position, which may yield a series of disagreements among group members and thereby break the chain of reasoning.

In contrast, teachers' argument development moves had a positive effect on students' production of causal chains and the effect lasted for three speaking turns. In Collaborative Reasoning discussions, teachers are encouraged to support student autonomy and provide minimal assistance. Consequently, teachers only intervene when necessary; for example, in situations such as students fighting for speaking turns or going off topic. In other words, teachers' moves are usually triggered by students' poor management of the discussion. However, students' leadership moves are more spontaneous and diverse; for example, a turn management move is made when a child wants to know how others think about the central question, an argument development move is made when a child does not understand the logic behind the speaker's argument, and a clarifying and revoicing move is made when the previous speaker's argument causes misunderstanding or confusion.

The acceptance or rejection of a leadership move affected whether a chain of reasoning was extended. Teachers' moves were seldom rejected by students; however, students' leadership moves were accepted or rejected conditioned on many factors such as cognitive ability and social centrality. If a student's leadership move was accepted, it is more likely for the group to develop a chain of reasoning. However, rejection of leadership moves may result in conflict among students. Under such circumstances, students may go off topic or the teacher may have to intervene to resolve the conflict, either of which would break the chain of reasoning. Below, I quote four excerpts from the discussion transcripts to further illustrate the different effects of peer-initiated and teacher-initiated moves on children's causal reasoning.

The first excerpt illustrated how an emergent leader, Rafael, prompted group members to contribute their ideas and eventually led to the production of a long multilink causal chain. As shown in the excerpt, Rafael first invited Olivia to give her opinion on whether or not wolves should be killed. Olivia provided her position and produced a chain of reasoning based on the interdependent relationships among the wolf, the elk, and the plants. Emilio (a quiet child) asked a clarifying question and Rafael resolved his confusion. Rafael's clarification also prompted Olivia to revise her speech. Building on Olivia's thinking, Rafael, Olivia, Tracy and Akira produced three different branches to the wolf-elk-plants chain. Rafael argued that the decrease of plants would affect the number of cows and people would not be able to eat steak if cows were gone. Tracy and Akira argued that the decrease of plants would cause the decrease of oxygen and people need oxygen to live. Olivia argued that vegetables were also plants, so people would starve if there were no vegetables to eat. Later, Rafael invited the quiet student, Emilio, to share his ideas. However, Olivia took the floor and repeated her previous argument. Rafael stopped Olivia and gave the floor back to Emilio. Emilio then shared his thoughts with the group.

Rafael	**Olivia, would you like to talk?**	[turn management]
Olivia	Yes, um…um…I think they should kill the wolves because um the wolves can keep on killin' the elk so we can have more plants, because the elk eat the plants that we do that um the plants.	
Emilio	**What plants?**	[Clarifying and revoicing]
Rafael	**Oh the producers**.	[Clarifying and revoicing]
Olivia	Eat the producers. That the producers are plants and the elk eat the plants so I think they should shouldn't kill the wolves 'cause they keep on killing the elks so we have more plants.	
Rafael	I'm with her 'cause like cows eat cows eat plant and we need steak and -	
Tracy	Oxygen.	
Olivia	We need vegetables and stuff.	
Tracy	We need oxygen.	
Rafael	Well like it can grow back if they eat the trees 'cause it rains.	
Akira	But you are gonna need like seeds and stuff.	
Olivia	'Cause if the wolves are gone, if the wolves are gone, the elk is gonna keep on eating the um um plants because there is nobody to kill them and then we're not gonna have enough because the elk is eating the plants. Not enough plants.	
Akira	We need oxygen. They have plants they have food.	
Rafael	**Have you thought something yet? Emilio, have you thought something yet?**	[turn management]
Olivia	We do eat plants we are not going anybody. We eat stuff that grow from trees.	
Emilio	Um I think no.	
Rafeal	[talk to Olivia] **Shhh-Shhh-Shhh-Shhh**.	[turn management]
Emilio	'cause they 'cause they they gotta eat just like we gotta eat. I'm done.	

The above example showed that acceptance of peer's leadership facilitated the exchange of ideas among group members and consequently increased the possibility of producing causal chains. However, rejection of leadership moves broke the chain of reasoning. In the second excerpt, Rafael was still the emergent leader. He took five speaking turns to invite Emilio, Olivia, Jade, Alisha, and Juan to express their opinions. Emilio accepted Rafael's invitation and expressed his position. However, when Rafael and Tracy asked Emilio to justify his position Emilio refused to continue the conversation. Rafael did not insist that Emilio explain, but started to invite others to speak. Jade gave a short response, while Olivia and Alisha both ignored Rafael's invitation. Finally, Rafael invited Juan to express his opinion. Juan made an unclear statement about eagles. Both Tracy and Gabriel expressed their confusion. However, the challenge from peers did not motivate Juan to provide an explanation. Only when the teacher intervened, Juan elaborated his thoughts and produced a causal chain.

Jack	Let Emilio say something?	[Turn management]
Rafael	He's talkin' All right Emi- ... Emilio.	[Turn management]
Emilio	I, think they shouldn't, die.	
Rafael	Ok. Why?	[Argument development by student]
Tracy	Why?	[Argument development by student]
Emilio	Would you just go?	
Rafael	Well you know... Olivia?	[Turn management]
Akira	She's gonna look at the poster.	
Rafael	Jade.	[Turn management]
Jade	I think they shouldn't die ... um ...no.	
Rafael	Alisha?	[Turn management]
Tracy	She already said something.	
Rafael	Oh. Juan.	[Turn management]
Juan	<u>I think that they shouldn't 'cause we can have more eagles.</u>	
Tracy	You can't. It's illegal to hunt eagles.	
Gabriel	What? They eat eagles?	[Clarifying and revoicing]
Juan	<u>You can have more eagles 'cuz the elk.</u>	
Teacher	Explain Juan.	[Argument development by teacher]
Juan	<u>Because the wolves kills the elk, and then the eagle comes and eat them up. And if the wolves die, your um- the elk will be still alive and the- um eagle can't come.</u>	

When a leadership move was rejected, the student who initiated this move tended not to take further actions to prompt for a response. The student leader either moved on to a different person or relied on the teacher to intervene. The third excerpt showed how the teacher jumped in and made a turn management move when two students fought for speaking turns. In this excerpt, Demisha first gave her position on the wolf question. Dallas then asked Demisha to justify her position. In responding to Dallas's argument development move, Demisha provided a reason for why she thought people should kill wolves. Eric challenged Demisha's argument in the subsequent turn. Demisha did not like being challenged by Eric, so she asked Eric to stop talking. Eric fought back and asked Demisha to quit the floor. The teacher had to intervene and reminded them of the turn-taking rules. Following teacher's request, Eric and Demisha stopped fighting and gave the floor to Dallas.

Demisha	They should.	
Dallas	Why?	[argument development]
Demisha	Because it's destroy the town and kill other animals.	
Eric	<u>But then 'cause if they do there're gonna be more elk basically, More elk and more stuff.</u>	
Demisha	**Shut up Eric.**	[turn management by student]
Eric	**I start first. Be quiet.**	[turn management by student]
Teacher	**One at a time please.**	[turn management by teacher]
Dallas	<u>If it's more elk and it's gonna it's like less trees and if it's less trees is less oxygen and if it's less oxygen it's less people.</u>	

In the above excerpt, teacher's intervention played a key role in extending the chain of reasoning. If the teacher had not intervened, Demisha and Eric might take the subsequent turns to fight for the opportunity to speak. Conflicts among group members often lead to an instant change in group dynamics. Teachers' role under such circumstances is to maintain quality interaction among students. Therefore, teachers are more likely to continue prompting even when their leadership moves are ignored by students. The fourth excerpt demonstrates this point.

In this example, five students and the teacher co-constructed a supply model. During the discussion, Lucas first stated that wolves *should not* be killed because "wolves eat elk" was just like "we eat burgers." The teacher built on Lucas's thinking but changed the direction of reasoning. She pointed out that people would not be able to eat hamburgers if wolves killed rancher's cows, which suggested that wolves *should* be killed. The two opposing viewpoints were repeated several times by the supporters of each side.

After the teacher brought up an opposing viewpoint, Kate, Ryan, and Tina altogether challenged the teacher's argument because they thought this idea lacked supporting evidence. This was the first disagreement in the discussion. To resolve this cognitive conflict, the teacher asked the students to consider this idea as hypothetical rather than factual. Lucas, Kate and Ryan then accepted the idea that wolves could hurt ranchers' business by eating their livestock, but they argued that ranchers themselves were responsible for the loss because they ran animals near the forest. This is the second disagreement in this discussion. To resolve this cognitive conflict, the teacher asked students to think from the rancher's perspective and imagine that the rancher got mad and wanted to kill all the wolves. Evelyn and Ryan finally accepted the teacher's argument and started to extend this line of reasoning. Evelyn contributed two links to the causal chain by arguing that if the wolves killed ranchers' cows, ranchers would not be able to sell beef to the Happy Burger restaurant, and in turn they would lose money and would not be able to buy more cattle. Finally, Evelyn incorporated the two opposing viewpoints and came to a conclusion that there were both advantages and disadvantages of killing wolves.

Lucas	We eat burgers, McDonalds and stuff like that and all they eat is I don't think they had I don't think they eat that much. I think one thing I know for sure they eat is elk.
Kate	They eat when I was reading that they eat elk and moose, the Yellowstone bison.
Lucas	OK. Yeah we eat way more than they eat. They should they should be able to eat. something. We eat happy burger. We eat well burgers. We eat McDonalds. We eat chicken, we eat turkey like Thanksgiving just passed. We eat all that meat and they probably didn't they don't celebrate Thanksgiving and Christmas and stuff like that.
Teacher	**So what would you say to a rancher who said you know what that's one less burger you're gonna have because a wolf got my cow last night.**
Kate	How do they know it's a wolf?
Ryan	Oh well we can get resource from another state.
Lucas	But we won't be able to eat. We need to eat to survive and like if-
Ryan	That's what I just said.
Kate	But how did they know it was a wolf?
Teacher	**He saw with his own eyes.**
Tina	They didn't say that.
Ryan	He said he saw something that looked like a wolf.
Teacher	**I'm giving you a for instance what if that happen? What would you say to that rancher?**
Lucas	I'll be kind of mad because it's it's it's it's a wolf. It's gonna eat. We eat. They eat.
Teacher	**But it ate his cow.**
Kate	A cow. That's how they gonna survive.
Ryan	Well, that's technically his fault to be so close to the forest.
Teacher	**That's one less hamburger you guys get. That's about 100 less dollars he's gonna get.**
Ryan	No. More of loss for for him than us.
Teacher	**I bet he's mad I bet he wants every wolf shot in the area. You're gonna say to them.**

Evelyn	Because the farmers because happy burger Happy burger buys their beef.
Ryan	Because it's his stuff?
Teacher	**That's right.**
Evelyn	And if you don't give him money he can't buy more cattle.
Teacher	**That's right.**
Evelyn	So that's why I was thinking that they should and they shouldn't at the same time.

This excerpt indicated how the teacher prompted students to consider the other side of the issue and to develop a more encompassing situation model after resolving disputed ideas. A notable point in this excerpt is that the 'wolves eat animals for living' argument was repeated four times in the discussion—three times by Lucas and one time by Kate. And the 'wolves eat livestock" argument was repeated three times by the teacher. In a group discussion, repetition is often considered as "a discourse-cohesive device" (Johnstone, 1987, p. 207). Revoicing a previous statement shows active listening and participation and it also reminds the group of the importance of an argument. If one speaker repeats another speaker's statement, it indicates that the two speakers have a shared understanding of the problem. When the discussion becomes long, information starts to drop out of memory (Kirschner, Paas, & Kirschner, 2009). If the previous ideas or viewpoints are reiterated, speakers are more able to continue contributing new ideas. Future analyses are needed to verify the influence of repeating ideas or arguments on the development of multilink causal reasoning.

Fourth, this study found that multilink causal chains were less likely to appear when students focused on the evaluation of evidence or the discussion of alternate solutions. The multilink causal reasoning patterns identified in this study are concerned with whether or not people should be allowed to kill wolves. However, the alternate solutions proposed by children in some groups do not involve killing the wolves. These are solutions that the children would like to consider *after* they have decided that they do not want to kill the wolves but would still like to offer some protection to the community. Likewise, the discussion on whether wolves really threaten the townspeople is another way to avoid making a decision on killing or not killing wolves. These are questions that the children would like to consider *before* they make a decision.

Evaluating the evidence of wolves' threat tended to trigger a debate among group members about whether or not they should consider the information as supporting evidence for killing wolves. In the following example, Alondra quoted information from the text and stated that wolves killed a dog, which suggested that wolves should be killed. Monique then questioned if the dog was killed by the wolf because there was no solid evidence to prove this happened. Alondra restated the information in the text to defend her position, but Monique pointed out that it was a speculative statement in the text. The group split into two groups to argue about whether the reported wolf attack really happened. Eva supported Monique's argument and provided an alternative explanation for the dog's death, while Carlos and Alondra challenged Monique's argument by proposing hypothetical witnesses for wolf attack. The debate continued until Alondra changed the subject.

Alondra	Look Monique, it says on the um text that you know that you say um that says that dogs are just like same as dogs. Wolves are just same as dogs but it says that um in the Winona's messenger, it said that um this dog named Elmo was killed by wolves 'cause he was um he was protecting um-
Monique	But what if that wasn't a wolf. You they don't-
Alondra	That was the wolf, it says.
Monique	It says possibly a wolf attack they don't say it was a wolf attack.
Eva	Yeah cause they didn't really know if there is a wolf. Could have been a bear.
Carlos	Bear?
Monique	Yeah. 'Cause they live bears live in the forest like all different kinds of animals.
Carlos	What if if it [wolf] appeared, it's just it wasn't in the forest and it came.
Monique	Or or probably kill the dog you know it probably it probably been bears 'cause only wolves and ecosystem it says right here it says the bears is in the forest, too.
Alondra	The dog was protecting a sheep because the wolf um wolves wolves.
Lola	He was trying to protect the sheep from a pack of wolves.
Monique	I know but how they don't actually know was if it was a wolf or not, so you can't assume that was.
Carlos	Well maybe they they call the bio* biologist.
Alondra	Maybe maybe the sh* maybe the sheep um run away.

Carlos	Hit hit wolves. That's how they want to kill each other.
Alondra	Maybe the sheep went running for the farmer and then they saw pack of wolves killing the dog.
Carlos	Well you don't know either if they saw the pack of wolves or not.
Monique	I know but you don't you can't you can't assume that that wolves killed killed um killed Elmo.
Carlos	See they even reported howling wolves.
Alondra	I think they should not get professional hunters because a hunter um lose lose toe in the accident.

This above excerpt indicated that the evaluation of evidence caused disagreement among group members and thus broke the chain of reasoning. The group split into two factions to support each side of the argument and took the following speaking turns to argue back and forth until one faction was persuaded by the other or someone brought up a new topic. The disagreement usually took many speaking turns to resolve and sometimes the group had to drop the topic to maintain rapport in the group. Although this form of arguing has its own value in a discussion, it takes up space that could be used on the development of a multilink causal chain.

Similarly, when someone brought up an alternate solution, others tended to start evaluating these solutions and the group engaged in long conversations about what kinds of fence they should build, how to capture the wolves, where to settle the wolves, how to train wolves, or how to transport local people to a different place. In the following example, Gabriel first proposed that they could domesticate wolves to avoid killing them. Rafael raised the question of how to capture the wolves. Gabriel proposed a plan and discussed with Rafael about possible problems that they might encounter and how to solve these problems. Tracy and Akira challenged the domestication idea. Gabriel proposed another alternate solution in response to Tracy and Akira's challenge, which was to build a fence to separate wolves and people. Rafael continued raising questions about this plan. Gabriel revised his plan in response to Rafael's challenges and elaborated on what kinds of fence they should build.

Gabriel	I think no [wolves should not be killed] because they can lock them up in a cage and feed them every once in a while when they are supposed to be fed.
Rafael	But how do they go get them in a cage 'cause you know the wol* the wolves.

Gabriel	Bring meat and-
Rafael	But like OK but the hunters the hunters can do that though. That's what I'm talking about.
Gabriel	That's what I'm talking about like, they can bring meat and tra* like put them on the ground so they can follow it, keep eating it and get in- they put some in the cage and then they go in there and they get locked up in there. Like in the movie where that have that big cage.
Rafael	But the wolves will know The wolves will know there is a cage like on the movie.
Gabriel	Yeah like that ca* cage in the movie.
Rafael	They'll know they locked up in something.
Gabriel	But those like like tall.
Tracy	Those steel types.
Rafael	Yeah, but they already but they'll know they are about to go in something. And they can sense, like they can smell, they probably they probably might see hunter with the gun.
Akira	And they'll get right back out because they, um, there's a hole in the cage.
Rafael	And right and no no. They can dig their way out.
Gabriel	No, I'm talkin' about like uh- I'm talkin' about like a cage, but then like a fence fence, like round fence, wooden fences.
Rafael	They can break dat fence. They can hop over it.
Gabriel	I'm talkin' about a wooden fence but it's got sumpin' really strong behind it, on the outside of it.

As shown in the above excerpt, alternate solutions stimulated the discussion of topics that had little connection with the central question and took up the time that could have been used on developing chains of reasoning addressing whether the wolves should be killed. Children sometimes produced chains of reasoning when discussing alternate solutions to the wolf question, but these chains were not evaluated in the present study. Although some alternate solutions have a level of endorsement by policy makers and wolf biologists, these ideas suffer from dubious premises that make them less acceptable as compared to patterns of multilink causal reasoning that can be justified based on the information in the Wolf Unit.

The detailed examination of social dynamics at work during collaborative discussions reveals that students spontaneously adopt a variety of roles related to their social and cognitive characteristics. For example, the "doers" make plans and try out different solutions to the problem, the "educators" explain to less skilled learners and summarize for less involved group members, and the "facilitators" invite quiet group member to speak and resolve conflicts among group members (see Brown & Palincsar, 1986; Chiu, 2000). The reason that students take different roles during collaborative group work is that they are accountable for their participation and for the way in which they respond to others. The mutual expectation between speakers and listeners lays the foundation for the development of shared understanding. The speaker is expected to present a clear and sound argument to the group, while the listener is expected to evaluate the argument and extend the line of reasoning. If a child is hesitant to voice his or her opinion, other children in the team are responsible for making room for the child to speak. If a child produces an ill-developed argument, group members are expected to ask clarifying questions to fix confusion and incoherence. If a child presents a divergent opinion, group members are anticipated to ask follow-up questions to understand its logic. And when disputes arise, the group is expected to weigh reasons and evidence and reach a resolution.

5.3 Some Limitations of this Study

This study has several limitations. First, the causal analysis of multilink chains in the wolf decision letter explains the variation between collaborative groups, as a function of frequency of self-initiated and peer-initiated chain construction during the summative Collaborative Reasoning discussion of the wolf question, but does not explain the difference between students who participated in collaborative group work (CG) and students who received direct instruction (DI). My previous study (Ma & Anderson, 2015) compared CG students' production of multilink causal chains to DI students' chain production in wolf decision letters and indicated a significant advantage for CG students. However, this condition difference has not been related to the social process in which DI students developed, or may have developed, multilink causal reasoning. The next step of this research is to transcribe the summative lesson in the DI condition and code multilink causal reasoning during the lesson to examine the possible mediating effect of teacher-guided chains of reasoning. The summative lesson in the DI condition was teacher-led final review of major concepts in the Wolf Unit. Future analysis will examine whether students' or teachers' chain production during the final review mediates the effect of direct instruction on students' chain construction in wolf letters, and then compare the mediating effect of teacher-guided chain construction to the mediating effect of peer-managed chain construction on individual chains of reasoning in the wolf decision letter.

A second limitation is that, while longer multilink causal reasoning chains were observed in the wolf decision letters of students who experienced six weeks of collaborative group work as compared to students who experienced six weeks of direct instruction, some may wonder whether this was due to an unmeasured pre-intervention advantage for the CG condition. The findings would have been more convincing if there had been an essay writing assessment before the intervention to determine if there were pre-intervention differences in causal reasoning. Please be clear that it is unlikely that pre-intervention differences exist, given the fact that this project was a well-implemented cluster-randomized clinical trial with control of a variety of factors including students' prior reading comprehension, oral English fluency, talkativeness, and home literacy resources and practices. The project does afford an indirect way to address possible pre-intervention differences. This is to transcribe and code multilink causal reasoning in the first CR discussion and look at whether or not the amount of multilink causal chains increases from the first to the second CR discussion and whether or not variation in multilink causal chain in the first CR discussion predicts students' chain production in the wolf letters. The suspicion of pre-intervention advantage would be eliminated if children produce fewer multilink causal chains in the first CR discussion and variation in multilink causal chains in the first CR discussion does not predict chain production in the wolf letters.

The third limitation of this study is that some of the hypotheses in earlier sections could have been evaluated in further analyses of this dataset to provide confirming or disconfirming evidence. For example, the interpretation of the finding that a causal link is more likely to follow the speaking turn of a child with poor reading comprehension is that group members may take the following speaking turn to explain ideas to this child, because this child's utterance was incomplete or confusing. A follow-up analysis would be to identify the speaking turns produced by poor readers, determine whether these turns tended to be incomplete or confusing, and examine if the subsequent speaking turns were devoted to explanation and clarification. Follow-up analyses of the present dataset are expected to provide evidence to support or refute hypotheses such as this one, but it was beyond the scope of this study to complete these analyses, as this would require further coding of the transcripts or the video.

Another limitation of this study is that we cannot properly evaluate the role of the Wolf Reintroduction and Management Unit in supporting the development of causal reasoning. The topic of wolf management affords the construction of long causal chains. Intuitively, students are likely to connect ideas to develop a chain of reasoning when the topic or knowledge domain, as realized in instructional materials, involves interconnected ideas and hierarchical relationships, but are less likely when topics do not involve a web of interconnections. For example, discussing a moral dilemma introduced in a short story may not result in multilink causal reasoning. Future studies are needed to examine the role of topic affordances in fostering construction of multilink causal chains.

5.4 Conclusions and Implications

In conclusion, this study demonstrates how causal reasoning is forged in talk and how peer collaboration facilitates individual cognitive growth through investigating children's construction of multilink causal reasoning chains during collaborative discussion and their application of causal reasoning in an independent essay. The study provides empirical evidence for internalization of a process of causal reasoning acquired during collaborative discussion and, thus, transfer of higher-order cognitive skills from one context to another. The positive effect of collaborative discussion supports the general claim that higher-order thinking and reasoning can be nurtured through fostering meaningful interaction among students.

This study demonstrates that student-student interaction and occasionally teacher-student interaction in proceeding speaking turns affect the likelihood that a causal chain is extended or suspended in subsequent speaking turns. The agreement among group members and support from leaders and socially centered students extended the chain of reasoning. However, refutation and disagreement stopped the chain because the group had to resolve disputed ideas in order to develop a shared understanding. Students' or teacher's redirection of topic served as resetting mechanism for the chaining process. Once the topic was reset by a student or by the teacher, it was more likely for the current speaker to start a causal chain.

The detailed analysis of the moment-by-moment exchange within groups of children provides a more complete and precise picture of the process from which an important new intellectual competency emerges, as compared to traditional instructional research that relies on end-of-unit or end-of-year tests. The microgenetic analysis opens a new frontier in child development research. Previous research on child development has largely focused on the change from one point to the next point, rather than the process in between the points in which higher-order thinking and reasoning are formed. Based on the findings of this study, developmental science can be hopeful about understanding major aspects of intellectual development by subjecting them to direct examination.

This study suggests that collaborative group work is an effective instructional approach for promoting children’s higher-order thinking and reasoning. The positive effect of collaborative group work raises our aspiration about what can be accomplished in schooling. It is important to note that participants of this study were predominantly Spanish-speaking English language learners or African American children from low-income homes. These children have limited opportunity to receive language support outside school and have limited access to literacy resources. While modern schooling places more and more emphasis on the development of higher-order thinking for ordinary children, the education of language learners and underrepresented minorities often focuses on basic academic skills such as decoding words or performing arithmetic operations. However, higher-order thinking and reasoning are not immediately achieved when students master basic academic skills. Instead, the ability to reason is forged in the day after day experience of exchanging ideas with others. Schools should pay more attention to major aspects of children's intellectual development in addition to basic academic skills, especially for underserved children. The major role of schooling in the information age is to provide an intellectually stimulating and personally engaging learning environment for students to acquire higher-order thinking and reasoning strategies, to develop social and communicative competence, and to learn how to manage their own learning process.

REFERENCES

Akkerman, S., Van den Bossche, P., Admiraal, W., Gijselaers, W., Segers, M., Simons, R. J., & Kirschner, P. (2007). Reconsidering group cognition: From conceptual confusion to a boundary area between cognitive and socio-cultural perspectives. *Educational Research Review*, *2*(1), 39-63.

Anderson, R. C., Chinn, C., Waggoner, M., & Nguyen-Jahiel, K. (1998). Intellectually stimulating story discussions. In J. Osborn & F. Lehr (Eds.), *Literacy for all: Issues in teaching and learning* (pp. 170–186). New York: Guilford.

Anderson, R. C., Nguyen-Jahiel, K., McNurlen, B., Archodidou, A., Kim, S.-Y., Reznitskaya, A., … Gilbert, L. (2001). The snowball phenomenon: Spread of ways of talking and ways of thinking across groups of children. *Cognition and Instruction, 19*, 1-46.

Aronson, E., & Patnoe, S. (1997). *The jigsaw classroom: Building cooperation in the classroom* (2nd ed.). New York, NY: Longman.

Assaraf, O. B. Z., & Orion, N. (2010). System thinking skills at the elementary school level. *Journal of Research in Science Teaching*, *47*(5), 540-563.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59*, 87-96.

Azmitia, M. (1996). Peer interactive minds: Developmental, theoretical, and methodological issues. In P. B. Baltes, & U. M. Staudinger (Eds.), *Interactive minds: Life-span perspectives on the social foundation of cognition* (pp. 133-162). Cambridge, UK: Cambridge University Press.

Barman, C.R., & Mayer, D.A. (1994). An analysis of high school students' concepts & textbook presentations of food chains & food webs. *The American Biology Teacher, 56*, 160-163.

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *The Journal of the Learning Sciences, 9*(4), 403-436.

Barron, B. (2003). When small groups fail. *Journal of the Learning Sciences, 12*, 307-359.

Barry, S., & Lazarte, A. A. (1998). Evidence for mental models: How do prior knowledge, syntactic complexity, and reading topic affect inference generation in a recall task for nonnative readers of Spanish?. *The Modern Language Journal, 82*(2), 176-193.

Bell, N., Grossen, M., & Perret-Clermont, A. N. (1985). Sociocognitive conflict and intellectual growth. *New Directions for Child and Adolescent Development*, *1985*(29), 41-54.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written communication*. Hillsdale, NJ: Erlbaum.

Black, S. D., Levin, J. A., Mehan, H., & Quinn, C. N. (1983). Real and non-real time interaction: Unraveling multiple threads of discourse*. *Discourse Processes*, *6*(1), 59-75.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCInet for Windows: Software for Social Network Analysis*, Harvard, MA: Analytic Technologies.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education*, *24*. 61-100.

Brown, A., & Palincsar, A. (1986). Guided cooperative learning and individual knowledge acquisition (Tech. Rep. No. 372). Urbana-Champagne: University of Illinois, Center for the Study of Reading.

Butts, C. T. (2008). Social Network Analysis with sna. *Journal of Statistical Software, 24,* 1-51. Retrieved from http://www.jstatsoft.org/v24/i06/

Cain, K., & Oakhill, J. V. (1999). Inference making and its relation to comprehension failure. *Reading & Writing*, *11*, 489-503.

Callanan, M. A., & Oakes, L. M. (1992). Preschoolers' questions and parents' explanations: Causal thinking in everyday activity. *Cognitive Development*, *7*(2), 213-233.

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.

Chi, M. T. (2006). Two approaches to the study of experts' characteristics. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 21-30). Cambridge, UK: Cambridge University Press.

Chi, M. T., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist*, *47*(3), 177-188.

Chinn, C. A., Anderson, R. C., & Waggoner, M. A. (2001). Patterns of discourse in two kinds of literature discussion. *Reading Research Quarterly*, *36*(4), 378-411.

Chinn, C. A. & Clark, D. B. (2013). Learning through collaborative argumentation. In C. E. Hmelo-Silver, C. A. Chinn, C. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 314-332). New York, NY: Routledge.

Chiu, M. M. (2000). Group Problem-Solving Processes: Social Interactions and Individual Actions. *Journal for the Theory of Social Behaviour*, *30*(1), 26-49.

Chiu, M. M. (2008). Flowing toward correct contributions during groups' mathematics problem solving: A statistical discourse analysis. *Journal of the Learning Sciences, 17*, 415-463.

Chiu, M. M., & Khoo, L. (2005). A new method for analyzing sequential processes dynamic multilevel analysis. *Small Group Research*, *36*(5), 600-631.

Clark, A., Anderson, R. C., Kuo, L., Kim, I. H., Archodidou, A., & Nguyen-Jahiel. K. (2003). Collaborative reasoning: Expanding ways for children to talk and think in school. *Educational Psychology Review*, *15*, 181-198.

Day, S. B., & Goldstone, R. L. (2012). The import of knowledge export: Connecting findings and theories of transfer of learning. *Educational Psychologist*, *47*(3), 153-176.

Dillenbourg, P., Baker, M., Blaye, A. & O'Malley, C. (1995) The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds.), *Learning in Humans and Machine: Towards an interdisciplinary learning science*. (pp. 189-211). Oxford: Elsevier.

Dobao, A. F. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, *18*(4), 497-520.

Dong, T., Anderson, R. C., Kim, I. H., & Li, Y. (2008). Collaborative reasoning in China and Korea. *Reading Research Quarterly*, *43*, 400-424.

Ellis, R. (2005). *Instructed second language acquisition: a literature review*. Wellington, New Zealand: Ministry of Education.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological review*, *102*(2), 211.

Evagorou, M., & Osborne, J. (2013). Exploring young students' collaborative argumentation within a socioscientific issue. *Journal of Research in Science Teaching, 50*(2), 209-237.

Fawcett, L. M., & Garton, A. F. (2005). The effect of peer collaboration on children's problem-solving ability. *British Journal of Educational Psychology, 75*(2), 157-169.

Flood, R. L. (2010). The relationship of 'systems thinking'to action research.*Systemic Practice and Action Research*, *23*(4), 269-284.

Galton, M., Hargreaves, L., & Pell, T. (2009). Group work and whole-class teaching with 11-to 14-year-olds compared. *Cambridge Journal of Education*, *39*(1), 119-140.

Garton, A. F., & Pratt, C. (2001). Peer assistance in children's problem solving. *British Journal of Developmental Psychology, 19*(2), 307-318.

Glymour, C. (2001). *The mind's arrows: Bayes nets and graphical causal models in psychology*. Cambridge, MA: The MIT Press.

Golbeck, S. L. (1998). Peer collaboration and children's representation of the horizontal surface of liquid. *Journal of Applied Developmental Psychology, 19*(4), 571-592.

Golbeck, S. L & El-Moslimany, H. (2013). Developmental approaches to collaborative learning: Toward collaborative knowledge creation. In C. E. Hmelo-Silver, C. A. Chinn, C. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 41-56). New York, NY: Routledge.

Goodwin, G. P., & Johnson-Laird, P. N. (2005). Reasoning about relations. *Psychological Review*, *112*(2), 468.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: causal maps and Bayes nets. *Psychological Review*, *111*(1), 3.

Gopnik, A., & Wellman, H. M. (2012). Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. *Psychological Bulletin, 138*(6), 1085.

Gressick, J., & Derry, S. J. (2010). Distributed leadership in online groups. *International Journal of Computer-Supported Collaborative Learning*, *5*(2), 211-236.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, *116*(4), 661.

Hakkarainen, K., Paavola, S., Kangas, K., & Seitamaa-Hakkarainen, P. (2013). Socio-cultural perspectives on collaborative learning: Towards collaborative knowledge creation. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. O'Donnell (Eds.), *The International Handbook of Collaborative Learning* (pp. 57-73). London, England: Taylor & Francis.

Hastie, R. (2016). Causal thinking in judgments. In G. Keren & G. Wu (Eds.), *The Blackwell handbook of judgment and decision making* (pp. 590-628). Oxford, England: Blackwell.

Hatano, G., & Inagaki, K. (1991). Sharing cognition through collective comprehension activity. In L. B. Resnick, J. Levine, & S. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 331–348). Washington, DC: American Psychological Association.

Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis*. New York: The Guilford Press.

Heilig, J. V. (2011). Understanding the interaction between high-stakes graduation tests and English learners. *Teachers College Record, 113*, 2633-2669.

Hickling, A. K., & Wellman, H. M. (2001). The emergence of children's causal explanations and theories: Evidence from everyday conversation. *Developmental Psychology*, *37*(5), 668-683.

Hmelo-Silver, C. E. & DeSimone, C. (2013). Problem-based learning: An instructional model of collaborative learning. In C. E. Hmelo-Silver, C. A. Chinn, C. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 370- 385). New York, NY: Routledge.

Hume, D. (1748/1999, T. L. Beauchamp, Ed.). *An enquiry concerning human understanding*. Oxford, UK: Oxford University Press, Oxford Philosophical Texts.

Hummel, J. E., & Holyoak, K. J. (2005). Relational reasoning in a neurally plausible cognitive architecture. *Current Directions in Psychological Science, 14*(3), 153-157.

Hummel, J. E., Licato, J., & Bringsjord, S. (2014). Analogy, explanation, and proof. *Frontiers in Human Neuroscience, 8*.

Inagaki, K., & Hatano, G. (2006). Young children's conception of the biological world. *Current Directions in Psychological Science*, *15*(4), 177-181.

Inqscribe digital media transcription software (Version 1.0) [Computer software]. (2005). Available from https://www.inqscribe.com

Jacobs, G. (1998). Cooperative learning or just grouping students: The difference makes a difference. In W. Renandya & G. Jacobs (Eds.), *Learners and language learning* (pp. 145-171). Singapore: SEAMEO.

Jadallah, M., Miller, B., Anderson, R. C., Nguyen-Jahiel, K., Archodidou, A., Zhang, J., & Grabow, K. (2009). Collaborative Reasoning about a science and public policy issue. In M. McKeown and L. Kucan (Eds.), *Bringing reading research to life: Essays in honor of Isabel L. Beck* (pp. 170-193). New York: Guilford Press.

Johnstone, B. (1987). Introduction: Perspectives on repetition. *Text*, 205-214.

Kant, I. (1997). *Critique of pure reason* (P. Guyer & A. W. Wood, Trans.). Cambridge, England: Cambridge University Press. (Original work published 1781)

Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology, 57*, 227.

Khemlani, S. S., Barbey, A. K., & Johnson-Laird, P. N. (2014). Causal reasoning with mental models. *Frontiers in Human Neuroscience*, *8* (849), 1-15.

Kim, I. H., Anderson, R. C., Miller, B., Jeong, J., & Swim, T. (2011). Influence of cultural norms and collaborative discussions on children's reflective essays. *Discourse Processes, 48*(7), 501-528.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge: Cambridge University Press.

Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes, 39*, 125-128.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and language*, *29*(2), 133-159.

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, *25*(2), 306-314.

Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning and Instruction*,*17*(6), 708-721.

Koslowski, B. & Masnick, A. (2010). Causal reasoning and explanation. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development,* 2nd ed. (pp. 377-398). Oxford: Wiley-Blackwell.

Kuhn, D., Shaw, V., & Felton, M. (1997). Effects of dyadic interaction on argumentative reasoning. *Cognition and Instruction, 15*, 287-315.

Ladd, G. W., Kochenderfer-Ladd, B., Visconti, K. J. & Ettekal, I. (2012). Classroom peer relations and children's social and scholastic development: Risk factors and resources. In A. M. Ryan & G. W. Ladd (Eds.), *Peer relationships and adjustment at school* (pp. 11-50). Charlotte, NC: Information Age Publishing.

Legare, C. H., Schult, C. A., Impola, M., & Souza, A. L. (2016). Young children revise explanations in response to new evidence. *Cognitive Development*, *39*, 45-56.

Li, Y., Anderson, R. C., Nguyen-Jahiel, K., Dong, T., Archodidou, A., Kim, I. H., ... & Miller, B. (2007). Emergent leadership in children's discussion groups. *Cognition and Instruction*, *25*(1), 1-2.

Lin, T.-J., Anderson, R. C., Hummel, J. E., Jadallah, M., Miller, B. W., Nguyen-Jahiel, K., … Dong, T. (2012). Children's use of analogy during Collaborative Reasoning. *Child Development, 83, 1429-1443.*

Lin, T. J., Anderson, R. C., Jadallah, M., Nguyen-Jahiel, K., Kim, I. H., Kuo, L. J., ... & Li, Y. (2015). Social influences on children's development of relational thinking during small-group discussions. *Contemporary Educational Psychology, 41*, 83-97.

Lin, T. J., Jadallah, M., Anderson, R. C., Baker, A. R., Nguyen-Jahiel, K., Kim, I. H., ... & Wu, X. (2015). Less is more: Teachers' influence during peer collaboration. *Journal of Educational Psychology*, *107*(2), 609.

Lin, T.-J., Ma, S., Zhang, J., Nguyen-Jahiel, K., Anderson, R. C., Morris, J. A., … Jadallah, M. (April, 2011). *Nurturing conceptual understanding and systems thinking.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Lu, J., Chiu, M. M., & Law, N. W. (2011). Collaborative argumentation and justifications: A statistical discourse analysis of online discussions. *Computers in Human Behavior*, *27*(2), 946-955.

Ma, S., & Anderson, R. C. (2015, April). *Transfer of multi-link causal reasoning in collaborative group work*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, Illinois.

Ma, S., Anderson, R. C., Lin, T. J., Zhang, J., Morris, J. A., Nguyen-Jahiel, K., ... & Yi, S. (2017). Instructional influences on English language learners' storytelling. *Learning and Instruction*, *49*, 64-80.

Ma, S., Lin, T. –J., & Anderson, R. C. (2013, July). *Nurturing English language learners' multi-link causal reasoning*. Paper presented at the 20th annual meeting of the Society for the Scientific Study of Reading, Hong Kong, China.

Ma, S., Zhang, J., Anderson, R. C., Morris, J., Nguyen-Jahiel, K. T., Miller, B., ... & Grabow, K. (2017). Children's Productive Use of Academic Vocabulary. *Discourse Processes*, *54*(1), 40-61.

MacGinitie, W. H., MacGinitie, R. K., Maria, K., & Dreyer, L.G. (2000). *Gates-MacGinitie Reading Tests (4th ed.), Level 4, Form S*. Itasca, IL: Riverside Publishing.

Mackey, A., & Goo, J. (2012). Interaction approach in second language acquisition. In C. A. Chapelle (Ed.), The encyclopedia of applied linguistics. Chichester, UK: Wiley-Blackwell.

Mason, L. (1996). Collaborative reasoning on self-generated analogies: conceptual growth in understanding scientific phenomena. *Educational Research and Evaluation, 2*(4), 309-350.

Mason, L., & Santi, M. (1998). Discussing the greenhouse effect: Children's collaborative discourse reasoning and conceptual change. *Environmental Education Research, 4*(1), 67-85.

Menken, K. (2009). No Child Left Behind and its effects on language policy. *Annual Review of Applied Linguistics*, 29, 103-117.

Mercer, N. (2008). The seeds of time: Why classroom dialogue needs a temporal analysis. *The Journal of the Learning Sciences*, *17*(1), 33-59.

Mercier, E. M., Higgins, S. E., & Da Costa, L. (2014). Different leaders: Emergent organizational and intellectual leadership in children's collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning*, *9*(4), 397-432.

Miller, B., Anderson, R. C., Li, Y., Nguyen-Jahiel. K., Kim, I. H., Jadallah, M., … Kuo, L. (2009). *Emerging Participation of Quiet Students in Collaborative Discussions.* Champaign, IL: Center for the Study of Reading.

Mogar, M. (1960). Children's causal reasoning about natural phenomena. *Child Development*, *31*. 59-65.

Mohammed, S., & Dumville, A. B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, *22*, 89–106.

Mohammed, S., Klimoski, R., & Rentsch, J. R. (2000). The measurement of team mental models: We have no shared schema. *Organizational Research Methods*, *3*(2), 123-165.

Morris, J., Anderson, R. C., and the Collaborative Reasoning Research Group. (2012). *The other side: Children's spontaneous generation of counterarguments*. Champaign, IL: Center for the Study of Reading.

Morris, J. A., Miller, B. M., Anderson, R. C., Lin, T.-J., Zhang, J., Sun, J., Scott, T., Latawiec, B., & Ma, S. (forthcoming) Instructional discourse and argumentative writing. *International Journal of Educational Research*.

Muthén, L. K., & Muthén, B. O. (2012). Mplus 7.4 [Computer software]. Los Angeles: Muthén & Muthén.

Nijstad, B. A., Stroebe, W., & Lodewijkx, H. F. (2002). Cognitive stimulation and interference in groups: Exposure effects in an idea generation task. *Journal of Experimental Social Psychology*, *38*(6), 535-544.

O'Donnell, A. M. & Hmelo-Silver, C. E. (2013). Introduction: What is Collaborative Learning? An overview. In C. E. Hmelo-Silver, C. A. Chinn, C. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp.1-15). New York, NY: Routledge.

Orsolini, M., & Pontecorvo, C. (1992). Children's talk in classroom discussions. *Cognition and instruction*, *9*(2), 113-136.

Osborne, J. (2010). Arguing to learn in science: The role of collaborative, critical discourse. *Science*, *328*(5977), 463-466.

Palincsar, A. S. (1998). Social constructivist perspectives on teaching and learning. *Annual Review of Psychology*, *49*(1), 345-375.

Park, J., & Sloman, S. A. (2013). Mechanistic beliefs determine adherence to the Markov property in causal reasoning. *Cognitive Psychology*, *67*(4), 186-216.

Perring, W., & Kintsch W. (1985). Propositional and situational representations of text. *Journal of Memory and Language, 24*, 503-518.

Phelps, E., & Damon, W. (1989). Problem solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology, 81*(4), 639.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Ploetzner, R., Dillenbourg, P., Preier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 103-121). Amsterdam, Netherlands: Pergamon.

Prentice, R. L., Williams, B. J., & Peterson, A. V. (1981). On the regression analysis of multivariate failure time data. *Biometrika, 68*(2), 373-379.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary educational psychology*, *23*(1), 90-108.

Reznitskaya, A., Anderson, R. C., & Kuo, L. J. (2007). Teaching and learning argumentation. *The Elementary School Journal, 107*(5), 449-472.

Reznitskaya, A., Anderson, R. C., McNurlen, B., Nguyen-Jahiel, K., Archodidou, A., & Kim, S. Y. (2001). Influence of oral discussion on written argument. *Discourse Processes, 32*(2-3), 155-175.

Reznitskaya, A., Kuo, L., Clark, A., Miller, B., Jadallah, M., Anderson, R. C., & Nguyen-Jahiel, K. (2009). Collaborative reasoning: A dialogic approach to group discussions. *Cambridge Journal of Education, 39*, 29-48. doi:10.1080/03057640802701952

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, *2*(3), 235-276.

Rosenblatt, L. M. (1989). Writing and reading: The transactional theory. In J. Mason (Ed.), *Reading and writing connections* (pp. 153-176). Boston: Allyn and Bacon.

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right... If they argue together!. *Cognition and instruction*, *18*(4), 461-494.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory, 6*, 174-215.

Sobel, D. M., & Legare, C. H. (2014). Causal learning in children. *Wiley Interdisciplinary Reviews: Cognitive Science*, *5*(4), 413-427.

Stahl, G. (2005). Group cognition in computer-assisted collaborative learning. *Journal of Computer Assisted Learning*, *21*(2), 79-90.

StataCorp. (2015). Stata Statistical Software: Release 14 [Computer software]. College Station, TX: StataCorp LP.

Stephenson, K., & Zelen, M. (1989). Rethinking centrality: Methods and examples. *Social Networks*, *11*(1), 1-37.

Storch, N. (2001). How collaborative is pair work? ESL tertiary students composing in pairs. *Language Teaching Research, 5*, 29-53.

Swain, M., Brooks, L., & Tocalli-Beller, A. (2002). 9. Peer-peer dialogue as a means of second language learning. *Annual Review of Applied Linguistics*, *22*, 171-185.

Sun, J., Anderson, R. C., Perry, M., & Lin, T-J. (2017). Emergent leadership in children's cooperative problem solving groups. *Cognition and Instruction*. doi: 10.1080/07370008.2017.1313615

Van den Bossche, P., Gijselaers, W., Segers, M., Woltjer, G., & Kirschner, P. (2011). Team learning: building shared mental models. *Instructional Science*,*39*(3), 283-301.

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, *24*(4), 535-585.

Vye, N. J., Goldman, S. R., Voss, J. F., Hmelo, C., & Williams, S. (1997). Complex mathematical problem solving by individuals and dyads. *Cognition and Instruction, 15*(4), 435-484.

Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard.

Waggoner, M., Chinn, C., Yi, H., & Anderson, R. C. (1995). Collaborative reasoning about stories. *Language Arts*, *72*(8), 582-589.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.

Webb, N. M. (2013). Information processing approaches to collaborative learning. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. O'Donnell. (Eds.), *The international handbook of collaborative learning* (pp. 19-40). London, England: Taylor & Francis.

Webb, N. M., & Mastergeorge, A. (2003). Promoting effective helping behavior in peer-directed groups. *International Journal of Educational Research, 39*(1), 73-97.

Wells, G. (1999) *Dialogic enquiry*. Cambridge: Cambridge University Press.

Wentzel, K. R., & Asher, S. R. (1995). The academic lives of neglected, rejected, popular, and controversial children. *Child Development*, *66*(3), 754-763.

Wentzel, K. R., Barry, C. M., & Caldwell, K. A. (2004). Friendships in Middle School: Influences on Motivation and School Adjustment. *Journal of Educational Psychology*, *96*(2), 195.

Xu, F. (2007). Rational statistical inference and cognitive development. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: foundations and the future* (Vol. 3, pp. 199–215). New York, NY: Oxford University Press.

Xu, F., & Denison, S. (2009). Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112*(1), 97-104.

Xu, F., & Kushnir, T. (2013). Infants are rational constructivist learners. *Current Directions in Psychological Science, 22*(1), 28-32.

Zhang, J., Anderson, R. C., & Nguyen-Jahiel, K. (2013). Language-rich discussions for English language learners. *International Journal of Educational Research*, *58*, 44-60.

Zhang, X., Anderson, R. C., Morris, J., Miller, B., Nguyen-Jahiel, K., Lin, T-J, … & Hsu, Y-L. (2016). Improving children's competence as decision makers contrasting effects of collaborative interaction and direct instruction. *American Educational Research Journal, 53*(1), 194-223.

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, *21*(3), 289-327.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological bulletin*, *123*(2), 162.